ANNALS of the
Association of American Geographers

Volume 48 JUNE 1958 Number 2

THE CONCEPT OF GEOGRAPHY AS A SCIENCE OF SPACE, FROM KANT AND HUMBOLDT TO HETTNER

RICHARD HARTSHORNE
University of Wisconsin

THE PROBLEM

NUMEROUS geographers writing in recent years concerning the nature and scope of their subject have described the relation of their field to other fields of science in terms of a concept said to stem from Immanuel Kant and from Alexander von Humboldt. Whatever may be the original source of the concept, its importance in current geographic thought stems from the writings of Alfred Hettner, the German master of the methodology of geography. We are not concerned here with the validity of the concept, which is of course in no way dependent on who originated or supported it.[1] As a study in the history of geographic thought, this paper is concerned with the possible origin, or origins, of the concept and its significance to geography during the past century and a half.

Hettner's first brief statement of the concept appears in his earliest methodological paper, with which he inaugurated in 1895 the *Geographische Zeitschrift*, the journal which he edited for forty years. Noting that the materials of study in geography included a vast diversity of facts, so that many had doubted whether they could be united in a single science, he wrote:

> If we compare the different sciences we will find that while in many of them the unity lies in the materials of study, in others it lies in the method of study. Geography belongs in the latter group; its unity is in its method. As history and historical geology consider the development of the human race or of the earth in terms of time, so geography proceeds from the viewpoint of spatial variations.[2]

Hettner published a full explanation of his concept a decade later, most completely in a paper analyzing the system of the sciences in the *Preussische Jahrbücher*,[3] somewhat less extended, as part of what was to become the most famous of his methodological papers, "Das Wesen und die Methoden der Geographie," in his own journal.[4] In contrast to the "systematic sciences" which study each a particular category of phenomena, whether of nature or of man, the historical or chronological sciences study the association of diverse phenomena in particular periods of time or in development through time, and the spatial, or chorological sciences study the associations of diverse phenomena in sections of space, or areas. In this sense, the historical

[1] An attempt was made to undermine the validity of the concept by challenging the legitimacy of its presumed origins, in Fred K. Schaeffer, "Exceptionalism in Geography," *Annals*, Association of American Geographers, Vol. 43 (1953), pp. 232–35. For detailed and documented demonstration of the errors and distortions on which that thesis is based, see Richard Hartshorne, "'Exceptionalism in Geography' Re-Examined," *Annals*, Association of American Geographers, Vol. 45 (1955), pp. 218–24. While Schaeffer's challenge was the stimulus which led to the present paper, this is not the study referred to in the statement "Preface to Two Papers," *Ibid.*, pp. 205–6. That second paper, which considers the validity of Hettner's concept, together with various other questions concerning the methodology of geography, is to be published shortly as a monograph in the new series to be published by the Association of American Geographers in cooperation with Rand, McNally and Co.

[2] Alfred Hettner, "Geographische Forschung und Bildung," *Geographische Zeitschrift*, Vol. 1 (1895), pp. 7–8.

[3] "Das System der Wissenschaften," *Preussische Jahrbücher*, Vol. 122 (1905), pp. 251–77.

[4] "Das Wesen und die Methoden der Geographie," *Geographische Zeitschrift*, Vol. 11 (1905), pp. 549–53.

sciences include historical geology, pre-history, and history proper (the history of literate peoples). The spatial sciences include astronomy and geography and, we may now add, geophysics. No sharp or absolute lines can be drawn between the three groups, for in many cases studies overlap, but the viewpoint is basically different in each case.

Immediately following the publication of the two articles in 1905, another German geographer, Schlüter, challenged the concept,[5] but Hettner was able to show that within the same paper Schlüter had expressed essentially the same conclusion.[6] In the following decades Hettner's concept became well known and widely accepted among German geographers. In his inaugural address at Edinburgh in 1908, Chisholm based his statement of geography on Hettner's concept.[7] In 1921 Michotte in Belgium based his orientation on Hettner's statement which he described as *"la classification habituelle."*[8] But in neither case does the concept appear to have been taken up by others in those countries. Even after the appearance in 1927 of Hettner's volume on the methodology of geography, which was widely acclaimed in other countries, little attention was paid to his re-statement of this basic concept.[9] The only student of the philosophy of science to give serious consideration to it, to the best of my knowledge, was Victor Kraft, in Vienna, who in 1929 discussed it and evidently found it in general acceptable.[10]

The concept has become widely known to English-speaking geographers as a result of its presentation, in 1939, in *The Nature of Geography*, essentially in Hettner's terms.[11] Numerous American and English geographers have used it as the basis for their consideration of the place of geography in the system of sciences.

Hettner evidently did not presume the concept was original with him. In his longer paper of 1905 he expressed surprise that the principal of the chorological sciences had escaped the attention of students who studied the classification of the sciences "even though a number of methodologists of geography have long declared it as the authoritative principle of geography."[12] He also noted that Kant had suggested this principle in his lectures on geography, but, according to a footnote to his other paper, this fact was brought to his attention only as he was completing his own writing—i.e., long after he had formulated his own concept.[13] In the republication in his volume of 1927 he introduced a quotation from Kant to demonstrate the similarity of ideas, but with no implication of any connection.[14] At no time does he appear to have recognized any connection between his concept and the views of Humboldt. The similarity between the two was first demonstrated by Döring in 1931,[15] and in 1939 I pointed out the similarity of ideas of all three students—Kant, Humboldt, and Hettner.[16]

Subsequent writers, using the materials presented in *The Nature of Geography*, have generally assumed that this demonstration of similarity established a direct connection.[17] But Hettner himself recognized no such connection and none has yet been established.

The purpose of the present paper therefore is to trace the history of the concept from its earliest origins to its exposition by Hettner in 1895 and 1905. Our concern is not merely with the bibliographical question but also with the more general problem of what conditions in the general climate of scientific thought may have caused students at certain

[5] Otto Schlüter, *Die Ziele der Geographie des Menschen* (Munich, 1906), pp. 53–6.

[6] "Methodische Streifzüge," *Geographische Zeitschrift*, Vol. 13 (1907), pp. 627–32.

[7] George C. Chisholm, "The Meaning and Scope of Geography," *Scottish Geographical Magazine*, Vol. 24 (1908), p. 567.

[8] P. Michotte, "L'Orientation nouvelle en géographie," *Bulletin de la Société Royale Belge de Géographie*, Vol. 45 (1921), p. 22.

[9] Hettner, *Die Geographie, ihre Geschichte, ihr Wesen und ihre Methoden* (Breslau, 1927), pp. 110–17.

[10] "Die Geographie als Wissenschaft," *Enzyklopädie der Erdkunde*, Teil: Methodenlehre der Geographie (Leipzig, Vienna, 1929), p. 8.

[11] Richard Hartshorne, *The Nature of Geography* (Lancaster, Pa., 1939, 1946), pp. 140–42.

[12] "Das System der Wissenschaften," *op. cit.*, p. 273.

[13] "Das Wesen und die Methoden der Geographie," *op. cit.*, p. 551.

[14] *Die Geographie . . . , op. cit.*, p. 115.

[15] Lothar Döring, *Wesen und Aufgaben der Geographie bei Alexander von Humboldt*, Frankfürter Geographische Hefte (1931).

[16] *The Nature of Geography*, pp. 134–35, 140–42.

[17] In several such studies, the reader might suppose that the author's conclusions were based on additional evidence because quotations from Kant, Humboldt, or Hettner are cited from the originals, whereas careful comparison of texts demonstrates that they were taken entirely from the intermediary source.

periods to overlook this concept whereas later students were to find it important to their thinking.

BEFORE 1750

Prior to the eighteenth century few students of geography felt any need to determine the status of their subject in the general field of knowledge; its importance was sufficiently assured by popular interest and general utility. In that century, however, an increasing number of students became concerned to establish geography as an integral field of knowledge, rather than merely a utility servant of commerce and government, or the handmaid of history. Geography, they were wont to assert, was similar and comparable to history—not a part of history, but coordinate with it.

This similarity has been recognized by so many students of many different countries that we may assume it to be readily observable as an empirical fact in the geographic literature. To seek its earliest origin we would no doubt need to go back to the period of ancient Greece, to men like Herodotus who wrote both history and geography.

The earliest definite statement of the comparison of geography and history that I have found is that of J. M. Franz in 1747.[18] At the same time, moreover, Franz and other students of that period recognized a close relationship between geography and astronomy, both of which they included under a common term—cosmography. Indeed the first geographical society in Germany, which Franz founded, was called *die Cosmographische Gesellschaft*.[19]

KANT AND HUMBOLDT—1756–1859

The earliest statement in which history and geography are not merely compared with each other but contrasted with the viewpoint of the systematic sciences, each defined in terms of categories of phenomena, is found in the introductory lecture of an elementary course in geography which Immanuel Kant gave at Königsberg during most of the second half of the eighteenth century. Kant himself never published these lectures, but numerous handwritten copies were circulated among students and at least one had been sent by Kant to a government official in Berlin.[20] More than a century later, Adickes found a score of such manuscript editions of Kant's course. Certain of these formed the basis for the publication by Rink, in 1802, of Kant's lectures.[21] By painstaking comparison of these many versions, Adickes has demonstrated that while the larger part of Rink's publication, the latter part, is based on materials used by Kant in the early years of the course and later discarded, the first part, including his statement of the relation of geography to other sciences, represents essentially the form in which Kant presented it in 1775 and subsequent years.[22]

Nearly a decade before the publication of Kant's lectures, the same basic concept of the nature of geography as a field of study was stated by Alexander von Humboldt in his first major publication, written in 1793 when he was 24 years old. Although Humboldt had been trained primarily in courses in economics and government finance in preparation for administrative work in government, his personal interest focussed on nature studies, particularly in botany and geology, and he had been introduced to geography in the field by George Forster, one of the first of the scientific explorers.[23] He evidently felt the need to establish a logical basis for a distinction between geography and other sciences. He out-

[18] Arthur Kühn, *Die Neugestaltung der deutschen Geographie im 18. Jahrhundert* (Leipzig, 1939), pp. 39, 41.

[19] *Ibid.*, p. 45 fn., pp. 54 ff.

[20] Erich Adickes, *Untersuchungen zu Kants physische Geographie* (Tübingen, 1911), pp. 3, 35 ff., 280 ff.

[21] *Immanuel Kant's physische Geographie*, edited by F. T. Rink (Königsberg, 1802). Although Kant had authorized this edition he had become too senile to examine what was included in it. An unauthorized version by Gottfried Vollmer was published in six volumes, beginning in 1801, but this has been shown to be based only in minor part on Kant's lectures. *Cf. The Nature of Geography*, pp. 38–39. Since Rink's edition is most readily available in various sets of Kant's collected works, where it appears with but minor changes in text but with varying page numbers, references to it in this paper are given by sections.

[22] Erich Adickes, *Ein neuaufgefundenes Kollegheft nach Kants Vorlesung über physische Geographie* (Tübingen, 1913), pp. 10–11, 67. The several paragraphs which form Kant's statement of the concept of geography under discussion in this paper are quoted in full in *The Nature of Geography*, pp. 134–35. In translating from Rink's edition, I incorporated corrections in the wording according to specific recommendations made by Adickes on the basis of his examination of the manuscripts, but this fact is not noted with the quotation, but in the previous footnote concerning Rink's edition, *Ibid.*, p. 39.

[23] *Ibid.*, pp. 49–50.

lined this distinction in a long footnote to his 1793 article and indicated subsequently that it continued to represent his concept of geography by re-publishing the same footnote in another article a decade later and again in the *Kosmos*, fifty years after its first publication.[24]

Since this statement of Humboldt's of 1793 is the earliest known publication of the concept under discussion, and because it is available only in the Latin form in which Humboldt published his first major work, it is translated here in full (from a photostat copy of the original, in the Library of Congress).[25] The more readily available re-publication in the Kosmos differs slightly in wording.

> Geognosy (*Erdkunde*)[26] studies animate and inanimate nature . . . both organic and inorganic bodies. It is divided into three parts: solid rock geography,[27] which Werner has industriously studied; zoological geography, whose foundations have been laid by Zimmerman; and the geography of plants, which our colleagues have left untouched. Observations of individual parts of trees or grass is by no means to be considered plant geography; rather plant geography traces the connections and relations by which all plants are bound together among themselves, designates in what lands they are found, in what atmospheric conditions they live, and tells of the destruction of rocks and stones by what primitive forms of the most powerful algae by what roots of trees, and describes the surface of the earth in which humus is prepared. This is what distinguishes geography from nature study,[28] falsely called nature history; zoology (*zoognosia*), botany (*phytognosia*) and geology (*oryctognosia*) all form parts of the study of nature, but they study only the forms, anatomy, processes, etc., of individual animals, plants, metallic things or fossils. Earth history, more closely affiliated with geography than with nature study, but as yet not attempted by any, studies the kinds of plants and animals that inhabited the primeval earth, their migrations and disappearance of most of them, the genesis of mountains, valleys, rock formations and ore veins . . . the earth surface gradually covered with humus and plants, denuded again by violent stream floods, and once more dried and covered by grass. Thus zoological history, the history of plants, and the history of rocks, which tell only the past state of the earth, are to be clearly distinguished from geography.

Both in substance and in terminology, this statement reflects the thinking of Abraham Gottlob Werner, under whom Humboldt was studying at the time, at the mining academy at Freiberg, Saxony. During the previous decades and more of teaching, Werner had separated materials formerly taught together in a single course into separate courses on minerals, which he called "*oryctognosie*," and on the study of rock formations and forms of mountains, which he called "*geognosie*" or "*Erdkunde*"—in either case, literally, "the knowledge of the earth."[29] Humboldt extended this distinction in respect to botany and zoology and also, apparently, added the comparison with the historical aspect of natural science.

There is no similarity in phrasing or in structure between this statement of Humboldt's and that of Kant. Nevertheless they are consistent in recognizing the same three divergent points of view in science. Was the statement which Humboldt published in 1793 inspired or influenced by the statement which Kant presented annually in his lectures since at least as early as 1775?

It is almost certain that there was no personal connection between the young Humboldt and the venerable Kant. During the period when they might have met, Kant never left Königsberg, and there is no mention of a visit by Humboldt to Königsberg in the voluminous correspondence which records his travels as a young man in Germany.

There were, however, many other ways in which Humboldt might have learned of Kant's concept. Before his college days, he and his older brother Wilhelm heard much of Kant's philosophy and work in physics in the intellectual circle in which they lived in Berlin.[30] His studies at the University of Frankfurt

[24] Alexander von Humboldt, *Kosmos: Entwurf einer physischen Weltbeschreibung*, Vol. 1 (Stuttgart, 1845), pp. 486–87.

[25] Alexander von Humboldt, *Florae Fribergensis Specimen* (Berlin, 1793), fn., pp. 9–10.

[26] In the original: "*Geognosia* (*Erdkunde*)"; both terms were evidently taken from Werner. In naming the three parts of "geognosy" in the next sentence, Humboldt used *Geographia* for each part. In the subsequent development of nomenclature of the sciences, geognosy came to be synonymous with geology or a part of it, whereas *Erdkunde* came to be synonymous with geography.

[27] "*Geographia oryctologica, quam simpliciter Geognosiam dicunt.*"

[28] "*Physiographia* (*Naturbeschreibung*) *historia naturalis perperam nuncupata.*" "Physiography" of course has come to have a very different meaning.

[29] "Abraham Gottlob Werner," in *Allgemeine deutsche Biographie*, Vol. 42, pp. 33–39; W. Blöde, "Die Geschichte und die jetzigen Verhältnisse der Bergakademie," in *Festschrift zum hundertjährigen Jubiläum der königlichen sächsischen Bergakademie zu Freiberg* (Dresden, 1866), p. 9.

[30] Karl Bruhns, ed., *Life of Alexander von Humboldt*, translated by J. and C. Lassell (London, 1873), Vol. I, p. 40.

made him familiar with Kant's philosophy. Wilhelm, with whom he was very close, was a great admirer of Kant, "has read all his works and lives and moves in his system." But Alexander himself appears to have had far less interest in the philosopher, even a negative reaction against him.[31]

It is, of course, possible that Humboldt could have seen one of the manuscript copies of Kant's lectures in geography, either at Frankfurt or at Göttingen, or in Berlin. But it seems unlikely that he would have been greatly interested in handwritten copies of elementary lectures—prepared and sold to enable students to pass a course—in comparison with the great amount of printed works that Kant had published.

In any case, there is no evidence that Kant's ideas about geography came to Humboldt's attention before 1793, or that he even was aware of Kant's interest in geography. In later years, however, Humboldt did almost certainly make use of the statement which Rink published for Kant in 1802.

Thus, in his lectures of 1827–28, if we can rely on the edition published a century later, he stated that his title "*physische Weltbeschreibung*" was taken from Kant.[32] But in explaining this title in his own publication, in the *Kosmos*, he merely says that it was an extension from the earth to the universe of "*die alte ausdrucksvolle Bennennung physische Erdbeschreibung*" and names no specific source.[33] This, however, is the term used in Rink's edition of Kant's lectures, whereas, as Adickes has shown, Kant probably said "*physische Geographie*" which Rink had changed to the Germanic form.[34] Likewise in defining his term "*physische Weltbeschreibung*," Humboldt wrote that it considers "*die Welt als Gegenstand des äusseren Sinnes*," placing that phrase in quotation marks but without reference;[35] the phrase is to be found in Kant's introductory lecture.[36] Further, as shown in the following passages, Humboldt contrasted his view of geography, or cosmology, with the "system of nature" of other sciences just as Kant had done, and with considerable similarity of phrasing—but again without reference to source.

> Humboldt: "Die systematisch geordneten Verzeichnisse aller organischen Gestaltungen, die wir ehemals mit dem allzu prunkvollen Namen von Natur-Systemen bezeichneten". . ."Die Einzelheiten . . . können logisch in Classen und Gattungen geordnet werden."[37]
>
> Kant: "Sage ich z.B. die Rindeart wird unter das Geschlecht . . . oder unter die Gattung . . . gezählt, so ist das eine Eintheilung, die ich in meinem Kopfe mache, also eine logische Eintheilung. Die *Systema naturae* ist gleichsam eine Registrar des Ganzen, wo ich alle Dinge, ein jedes in seine ihm eigenthümlich zukommende Classe setze."[38]
>
> Humboldt: "Solche Anordnungen führen . . . als ein naturbeschreibender Theil, den anmassenden Titel von Natur-Systemen . . . als Verzeichnisse gewähren sie nur ein formelles Band; sie bringen mehr Einheit in die Darstellung als in die Erkenntnis selbst."[39]
>
> Kant: "Indessen dürfte man die Systeme der Natur . . . richtiger wol Aggregate der Natur nennen, denn ein System setzt schon die Idee des Ganzen voraus, aus der die Mannigfaltigkeit der Dinge abgeleitet wird. Eigentlich haben wir noch gar kein *Systema naturae*. In den vorhandenen sogenannten Systemen der Art, sind die Dinge bloss Zusammengestellt, und an einander geordnet."[40]

We conclude, therefore, that (1) in later years Humboldt studied Kant's statement as published in 1802 and made important use of it; (2) when he wrote his own basic statement in 1793 he almost certainly did not have before him any of the numerous manuscript copies of Kant's lectures that were circulating in Germany; (3) it is entirely possible that he may at some earlier date have seen such a copy or may have heard of Kant's concept from any of many possible sources, but we have no scrap of evidence that he did. It is

[31] Rudolf Borch, *Alexander von Humboldt: Sein Leben in Selbstzeugnissen, Briefen und Berichten* (Berlin, 1948), pp. 26–27, 32–33, 41.

[32] Alexander von Humboldt, *Vorlesungen über physikalische Geographie nebst Prolegomenen über Stellung der Gestirne, Berlin im Winter 1827–28*, edited by Miron Goldstein (Berlin, 1934), p. 14.

[33] Humboldt, *Kosmos*, *op. cit.*, p. 52.

[34] *Kant*, *op. cit.*, Sec. 2; Adickes, *Ein neuaufgefundenes Kollegheft*, *op. cit.*, pp. 33–34. The term "physische" did not have for the contemporaries of Kant or Humboldt the meaning we now associate with "physical"—i.e., natural or exclusive of human. On the contrary both those students included under physical geography races, languages, and customs of man. The closest approximation of their concept of "physical geography" in present terms would be what Europeans call "general geography," Americans "systematic geography"; *cf.* Döring, *op. cit.*, pp. 15, 18; and *The Nature of Geography*, pp. 36, 43, 67, 76.

[35] *Kosmos*, *op. cit.*, p. 52.

[36] *Kant*, *op. cit.*, Sec. 2.

[37] *Kosmos*, *op. cit.*, pp. 55, 66.

[38] *Kant*, *op. cit.*, Sec. 4. According to Adickes, *op. cit.*, p. 35, Kant probably did not use the word "*Classe*," but that is what Humboldt would have read in Rink's edition of Kant's lecture.

[39] *Kosmos*, *op. cit.*, p. 66.

[40] *Kant*, *op. cit.*, Sec. 4.

entirely possible, if not probable, that the two men arrived at similar conclusions entirely independently.

OTHER GEOGRAPHERS IN THE FIRST HALF OF THE NINETEENTH CENTURY

It is difficult to demonstrate that either Kant's or Humboldt's statement had any significant influence on the thinking of other students of the time. The appearance of two conflicting editions of Kant's course at a time when he was too senile to judge either as authentic raised immediate doubts as to the reliability of either. The elementary form of the lectures was no doubt unimpressive and the greater part of their substantive material was clearly antiquated; not until a century later was it discovered, by Adickes' research, that for the latter part of the volume, the editor had used a manuscript of Kant's that was already over forty years old.

Humboldt's original statement, though published three times in all, appeared each time only as a footnote, and in Latin. The statements in his lectures of 1827–28 were not published for over a century, and his most detailed discussion is hidden in the midst of his long introduction to the *Kosmos*, where it is confused with several other questions he was endeavoring to clarify.[41]

In any case, most of the students of the time may have found the simple comparison with history adequate to assure status to geography. The term "science" had not yet become a fetish bestowing magical authority on those who acquired title to its use. Kant and Humboldt had both been attracted to geography from studies of nature rather than history. Each of them also had a universalist view of the field of knowledge and hence a concern to clarify the position of the subject he was presenting in relation to the total field. Few geographers then, or perhaps now, felt that need.

Carl Ritter evidently did not feel the need; so far as I can find he made no attempt to state the position of geography in relation to the whole field of knowledge. He did, however, express frequently the comparison of geography and history, and in one case at least in terms that seem to echo those of Kant.

Ritter: "Das Nebeneinander der Örtlichkeiten . . . das Nacheinander der Begebenheiten oder der Aufeinanderfolge und Entwicklung der Dinge."[42]

Kant: "Begebenheiten, die aufeinander folgen . . . Begebenheiten, die neben Einander im Raum vor sich gehen."[43]

Likewise in explaining the interest of geography in phenomena which are also the subject of study in the systematic sciences, Ritter's method of expressing the concern of the latter fields is similar to Humboldt's statement:

Ritter: "nach den Stoffen, Formen und inwohnenden Kräfte des materials an sich."[44]

Humboldt: "formas, anatomen, vires scutantur."[45]

Ritter also, it should be noted, recognized the logical similarity of astronomy and geography.[46]

In the methodological literature of the time of Humboldt and Ritter, and for half a century later, I have found but one publication that shows clear indication of familiarity with the statements which both Kant and Humboldt had made concerning the position of geography among the sciences. In a little-known essay published in 1834, Julius Fröbel stated the concept in terms that to me are clearer than those of Humboldt and in more complete form than those of Kant, for he recognizes geography as one of a group of "spatial sciences," the "cosmographic sciences."[47] Although the statement itself is presented without reference to sources, elsewhere in the same essay Fröbel quotes with references from other passages in Kant's introductory lecture and likewise from Humboldt's statement of 1793.[48] We may credit Fröbel therefore as the first writer—and so far as I can find, the only one prior to 1939—to have recognized the essential similarity of the statements of Kant and Humboldt. We know, however, from his autobiography that he had discussed geography personally with Humboldt a few years earlier in Berlin,[49] so that

[41] Pp. 48–73.

[42] "Über das historische Element in der geographischen Wissenschaft," in *Abhandlungen d.k. Akademie der Wissenschaft zu Berlin*, 1833, Hist.-philolog. Klasse, p. 41. (For other editions of this essay see the Bibliography in *The Nature of Geography*, pp. 4 and xxv.)

[43] *Kant*, *op. cit.*, Sec. 4.

[44] *Op. cit.*, p. 45.

[45] *Florae Fribergensis Specimen*, *loc. cit.*

[46] *Op. cit.*, p. 50.

[47] "Entwurf eines Systemes der geographischen Wissenschaften," *Mittheilungen aus dem Gebiete der theoretischen Erdkunde*, Vol. 1 (1834, published in Zurich, 1836), pp. 15–16.

[48] *Ibid.*, pp. 5, 12, 30, 123.

[49] *The Nature of Geography*, p. 73.

he may have learned of the similarity from Humboldt. In any case, he was the first to put the two statements together.

No one, however, appears to have noticed his statement. Fröbel remained in the profession only a few years during which he established something of a reputation as a promising but immature critic of methodology —particularly as a result of a published debate with Ritter.[50] Few readers, one may hazard, noted his statement of the concept hidden in the midst of nearly fifty pages of a new and elaborate structure for geography—an essay itself buried at birth by publication in a journal the author founded in Switzerland, which shortly died.[51]

Some forty years later, Hermann Wagner called attention to this long-forgotten essay, but only as an example of the type of study that could have no effect on the development of geographic thought; Wagner did not mention its statement of this concept.[52] Thereafter it appears to have been almost completely overlooked. The one copy which I have seen had rested in the Smithsonian Institute and the Library of Congress for over a century with its pages uncut.

IN THE SECOND HALF OF THE NINETEENTH CENTURY

Throughout the second half of the nineteenth century the statements of Kant and Humboldt on the place of geography among the sciences appear to have been completely overlooked.

In part this reflects the complete discontinuity in training of geographers at the university level following the death of Humboldt and Ritter, both in 1859. Humboldt never held a teaching position and no successor was appointed to Ritter's chair in geography. When professorships in geography were established in most German universities after 1871, they were filled by men who had not been trained by geographers.[53] For their understanding of the methodology of the field these new professors of geography were dependent on the published literature, particularly on the well-known essays of Carl Ritter. These were discussed at length by such students as Peschel, Marthe, and Ratzel.[54] But while Ritters' thinking, as we have noted, was consistent with the concept which Kant and Humboldt had stated, he had not directly expressed that concept in his own writings nor had he referred readers to the statements of Humboldt or Kant.

Humboldt's work was regarded as of great importance, but primarily for its descriptions of the countries he had visited. Students endeavored to induce his methodolgy from those writings but overlooked his scattered statements on methodology itself. Thus as late as 1927, Hettner stated, in his history of geography, that Humboldt had never concerned himself with the methodology of geography and that the famous work of his old age, the *Kosmos*, was not expressly a geographic work since it combined general geography (systematic geography) with astronomy.[55]

Fröbel's re-presentation of the concept, as we noted earlier, appears to have become well-nigh completely lost.

In view of the frequency of mention of Kant's name in modern discussion of the nature of geography, it may seem surprising that geographers of the nineteenth century paid so little attention to what he said on the subject. For this there were a number of reasons.

Students of the published works of Humboldt and Ritter would find no reason to look to the philosopher Kant for ideas about geography. If each of those masters, as we have suggested, used particular ideas and phrases from Kant in their writings, neither mentioned the source. The substantive materials of his lectures, as published in his name, offered nothing of value. Kant had also published a few individual research studies, on the origin of winds, of volcanism, etc., but neither Humboldt nor Ritter, so far as I have found, ever mentioned these.

Indeed there is strong negative evidence to indicate that both the founders of modern geography turned their backs on the geographic writings of the philosopher. Both were dominated in their thinking by the empirical approach to knowledge and distrusted the deductive thinking of *Natur Philosophie*.[56] On the basis of a detailed examination of the

[50] *Ibid.*, pp. 72–73, 102–6.
[51] *Ibid.*, p. 104.
[52] "Bericht über die Methodik der Erdkunde," *Geographisches Jahrbuch*, Vol. 7 (1878), pp. 621–22.
[53] *The Nature of Geography*, pp. 86, 106.
[54] *Ibid.*, p. 53.
[55] Hettner, *Die Geographie . . . , op. cit.*, p. 85.
[56] *The Nature of Geography*, pp. 54, 72; Bruhns, *Life of Alexander von Humboldt, op. cit.*, pp. 196–97, 205; Döring, *op. cit.*, pp. 27, 43–6.

many references to Kant in the *Kosmos*, Lind has demonstrated that while Humboldt spoke repeatedly of "the great philosopher," his concern was to attack Kant's scientific theories—frequently, Lind held, unfairly. Thus the astronomical theories which Kant had established or suggested on the basis of careful calculations, Humboldt described as having been "divined," "suspected," or "dreamed."[57]

The personal explanation which Lind suggests, namely that this is an example of the jealousy of one great man for another, near contemporary,[58] is mere conjecture and hardly plausible at that. The two were in fact not contemporaries, since Kant died before Humboldt became important. Further, Humboldt revealed no such characteristic in his relations with other men of high standing, as is shown particularly in his reflections on the work of Ritter.[59]

Nevertheless if we combine the facts which Lind presents of Humboldt's disparaging reflections on Kant's scientific work, with the fact that when he used Kant's ideas concerning geography, even in quotation marks, he did not mention Kant's name,[60] we can hardly escape the conclusion that Humboldt was motivated to belittle the scientific work of the philosopher. Why should he have done that?

In his earlier years, Humboldt had been strongly influenced by Goethe's philosophy of nature and at one time expressed interest in the system of the philosopher Schelling. But the subsequent development of a natural philosophy that would displace observation and experiment with pure reason and abstract ideas moved him to vitriolic condemnation of what he called a "mad saturnalia," a "*bal en masque* run mad."[61] He could most effectively undermine that school by disparaging the scientific accomplishments of the eminent philosopher.

Whatever the reasons, Kant's work and interest in geography was largely ignored for nearly a century after his death. In the very detailed histories of the development of geography before the nineteenth century which Peschel, Wisotzki, and Günther published in the latter part of that century, the relatively few references to Kant's studies in geography place him as but one of the scores of scholars who had made minor contributions.[62] Richthofen, writing in 1903, refers to Kant's work in philosophy and astronomy, but makes no mention of his work in geography.[63] Hettner's historical essay of 1898 does not mention Kant.[64] The fuller study of the "History of Geography" in his 1927 volume mentions Kant only as having preceded Laplace in presenting the nebular hypothesis of planetary origin and as having given a course in geography.[65]

It may well be that in the conflict between the empirical scientists and the *a priori* natural philosophers, the former were completely triumphant in the second part of the nineteenth century and Kant's very fame as a philosopher tended to obscure his scientific work. When later it was demonstrated that his studies in astronomical theory had represented contributions of outstanding importance,[66] several geographers were stimulated to examine his briefer studies in geography, but failed to find in them significant contri-

---

[57] P. von Lind, "Immanuel Kant und Alexander von Humboldt: eine Rechtfertigung Kants und eine historische Richtigstellung," *Zeitschrift für Philosophie und philosophische Kritik*, Vol. 106 (1895), pp. 68 ff., 256–57, 266–67, 270 ff., Vol. 107 (1896), pp. 28 ff. While Lind's numerous comparisons demonstrate in total that Humboldt gave less than justice to Kant's scientific work, Kaminsky warns that Lind was led by his "enthusiastic veneration" for Kant to numerous exaggerations unfair to Humboldt; see Willy Kaminski, *Ueber Immanuel Kants Schriften zur physischen Geographie* (Königsberg, 1905), p. 13.

[58] *Op. cit.*, Vol. 106 (1895), pp. 51–2.

[59] *The Nature of Geography*, pp. 53–54.

[60] That is, in no work which Humboldt himself published. In the edition of his Berlin lectures, published a century after they were given from notes taken by a listener, Kant is named as the source of his subtitle for cosmography and both editions of Kant's lecture course on geography are listed, but these may have been added by the listener or the editor; Humboldt, *Vorlesungen . . .*, op. cit.

[61] Humboldt, *Kosmos*, *op. cit.*, pp. 68–9; Bruhns, *Life of Alexander von Humboldt*, *op. cit.*, pp. 201–5.

[62] Oscar Peschel, *Geschichte der Erdkunde bis auf Alexander von Humboldt und Carl Ritter* (Munich, 1877); Emil Wisotzki, *Zeitströmungen in der Geographie* (Leipzig, 1897); Siegmund Günther, *Geschichte der Erdkunde* (Leipzig, 1904).

[63] Ferdinand Frh. von Richthofen, "Triebkräfte und Richtungen der Erdkunde im neunzehnten Jahrhundert," *Zeitschrift der Gesellschaft für Erdkunde, Berlin*, Vol. 38 (1903), pp. 672, 679.

[64] Alfred Hettner, "Die Entwicklung der Geographie im 19. Jahrhundert," *Geographische Zeitschrift*, Vol. 4 (1898), pp. 305–20.

[65] *Die Geographie . . .*, *op. cit.*, pp. 68, 71.

[66] W. Hastie, *Kant's Cosmology* (Glasgow, 1900), pp. xvii ff., xlv ff.

butions to substantive knowledge.[67] Hence they were not motivated to consider his statement of the position of geography among the sciences.

Nevertheless, even though the several statements of Kant and Humboldt, and that of Fröbel were not forced upon the attention of later geographers, they existed in the published literature. It is too much to assume that no geographer saw them; rather we must assume that anyone reading them failed to respond favorably. The reasons for this are to be found in the character of development of geography, and of science in general, in the second half of the nineteenth century.

As we noted earlier, when geography became permanently established in German universities, it was promoted largely by men trained in other fields—in a great variety of other fields. The consequence was methodological confusion—and vigorous methodological discussion.[68] In part also the confusion was forced on geography by certain characteristics in the general development of science in that period—notably the emphasis on the arbitrary separation of nature and man, which earlier students had not accepted, and the short-sighted view that the end-purpose of science was the construction of scientific laws.

The new generation of geographers, bringing these concepts from the fields in which they had been trained, produced a double form of dualism in geography. Physical geography—notably in the study of the origin and development of land forms—could claim a place as a natural science constructing and applying scientific laws. Human geography, in contrast, not only had drifted farther from its physical base, but in focussing on the study of particular areas could construct no laws and hence appeared unscientific. Ratzel demonstrated that this contrast was unnecessary by laying the foundations of systematic human geography, but there still remained the dualism between physical geography as a natural science and human geography conceived as a sort of missionary bridge from the natural sciences to the less securely founded social studies. As long as this dualistic viewpoint prevailed, there was no place for the concept Kant and Humboldt had formulated. A Gerland could read Kant's statement and pass it on, as he did in 1905, without seeing anything of value in it.[69]

These discussions, however, ultimately led to the replacement of the dualistic viewpoint by a unified orientation of geography. This was most effectively stated by Richthofen in his Leipzig inaugural address of 1883, which was recognized as restoring the viewpoint of the field that was common to the work of both Humboldt and Ritter and subsequently became widely accepted among German geographers as the programmatic statement of modern geography.[70]

Neither in this nor in any other paper did Richthofen refer to the concept of the place of geography among the sciences, but his discussion of the nature and scope of geography is consistent with that concept. We know that he, and Hettner who studied with him, read widely in Humboldt's substantive works. To what extent was either of them influenced by Humboldt's methodological viewpoint absorbed in such study? We can hardly hope for an answer to that question. More significant is the logical effect of the re-establishment of the earlier orientation. For the reassertion of geography as a unified integral field independent of the division that had become established between the natural and the social sciences inevitably raised the question of how geography could be fitted logically into the total system of knowledge, in a logical classification of the sciences.

HETTNER, 1895–1927

For Hettner, who had been trained in philosophy as well as in geography—and had even considered at one time going into philosophy—it was natural to seek an answer to

[67] Several of these studies of Kant's work are reviewed at some length in Kaminski's dissertation of 1905, *op. cit.*, pp. 6–21. In the same year the most thorough study of Kant's work in geography was published by Gerland, who concluded that Kant was less concerned to make positive contributions to geography, rather was concerned to establish conclusions in geography significant for his philosophy; George Gerland, "Immanuel Kant, seine geographischen und anthropologischen Arbeiten," *Kant-Studien*, Vol. 19 (1905), pp. 508 ff.

[68] Ernst Plewe, "Vom Wesen und den Methoden der regionalen Geographie," *Studium Generale*, Vol. 5 (1952), p. 411; Johann Sölch, "Die wissenschaftliche Aufgabe der heutigen Geographie," *Almanach der Oesterreichischen Akademie der Wissenschaften*, Vol. 98 (1948), pp. 146–47.

[69] *Op. cit.*, pp. 502–3.

[70] Solch, *op. cit.*, p. 147; Hettner, *Die Geographie . . . , op. cit.*, p. 106; *The Nature of Geography*, pp. 91, 116, 121, 136.

this question. He had studied with Richthofen both before and after the latter's inaugural address at Leipzig, and in between had done field work in South America.[71] While he accepted Richthofen's statement he found it provided no answer to this general question. He does not tell us the sources of his thinking but evidently he found in the concept of geography as a chorographic science, which Richthofen had taken over from Marthe and others, the essential comparison with history in contrast to the systematic sciences. Even before Richthofen's address, Wagner had distinguished between those geographers who found the basic concept of geography as an independent science in the "object" which it alone studied and those who found it in its "distinctive method of study."[72]

In the methodological discussion of the latter part of the century, the nearest approach to Hettner's concept that I have found is in an address by the Italian geographer, Dalla Vedova, published in 1881 and discussed the following year by Wagner in the *Geographisches Jahrbuch*, a discussion which we may assume Hettner probably read. The individual sciences study their objects from three points of view: the "static" viewpoint, according to the character of the phenomena at a given moment; the "dynamic" according to the manner of their existence and development in time; and the "chorological," according to their collective existence in space. The third viewpoint presents the field open to geography.[73] The original article, which is well documented, indicates no connection back to the statements of either Kant or Humboldt.[74] Neither does it resemble Hettner's later statement save in the recognition of the three points of view.

There is no reason to suppose that Hettner would ever have examined the obscure essay in which Fröbel restated the concept of Kant and Humboldt. The only mention of that writer that I have found in Hettner's writings is in a footnote stating that he excluded methodological views that had no effect on later developments, such as "the methodological demands of Fröbel[75]—presumably those in his debate with Ritter.[76]

According to his own statement, as noted earlier, Hettner in writing his basic statement in 1905 was unaware that Kant had recognized a similar viewpoint as the basis for recognition of geography as a separate science. Comparison of the two statements shows no similarity in organization or phrasing.[77]

Likewise Hettner appears not to have known that Humboldt had presented essentially the same concept. He would have had no reason to look in Humboldt's study of subterranean vegetation, published in Latin, in which the concept is stated in a long footnote. In the later discussion, in the *Kosmos*, Humboldt was concerned to establish a single science of cosmology, including both astronomy and general (systematic) geography, whereas Hettner considered astronomy and geography as separate sciences each concerned with a different section of space. Further, Humboldt's concept of the terrestrial portion of his cosmology included the entire earth-body, whereas Hettner followed the practice of Ritter and most geographers in limiting the scope of the field to the thin outer shell of the planet —"the earth surface." Finally, Humboldt's cosmology separated general or systematic geography, which he included in his cosmology, from special or regional geography, whereas Hettner, following Richthofen and most geographers since Varenius, included both in the single field of geography.

Hence, though Hettner believed his view of geography was in general consistent with that expressed in Humboldt's substantive works, we can believe his statement (made to me later in correspondence) that his formulation of the concept of the position of geography among the sciences was independent of such statements by Humboldt. Certainly there is no similarity either in phrasing or in the organization of his presentation.

Hettner's own view of his over-all contribution to the development of methodology in geography no doubt applies to this specific case: "My own importance in the construction

[71] Heinrich Schmitthenner, "Alfred Hettner," in Alfred Hettner, *Allgemeine Geographie des Menschen* (Stuttgart, 1947), Vol. 1, xi–xxxiv.

[72] Hermann Wagner, "Bericht über die Methodik der Erdkunde, *Geographisches Jahrbuch*, Vol. 9 (1882), p. 678.

[73] *Ibid.*, p. 680.

[74] G. Dalla Vedova, "Il Concetto Popolare e il Concetto Scientifico della Geografia," *Bulletino della Societa Geografica Italiana*, Vol. 18 (1881), pp. 5–27.

[75] Hettner, "Die Entwicklung der Geographie im 19. Jahrhundert," *op. cit.*, p. 305.

[76] See footnote 64.

[77] *The Nature of Geography*, pp. 134–35, 140–42.

of the methodology of geography has been exaggerated; I believe only that I have clearly expressed and methodologically established what was actually present in the development of the field."[78]

RE-DISCOVERY OF THE CONCEPTS OF KANT AND HUMBOLDT—1905–1939

The development of methodological thought in German geography toward the end of the nineteenth century no doubt contributed to the coincidence that Kant's long-forgotten statement was re-discovered just when Hettner was writing his basic statement of 1905. The observance of the hundredth anniversary of Kant's death caused Friedrich Hahn, professor of geography at Koenigsberg, who had earlier worked with Richthofen at Leipzig, to re-examine the geographic works of his famous predecessor. He sensed, and his student Kaminsky demonstrated in his doctoral dissertation, that Kant's importance in geography was not to be sought in his few substantive studies, but in his teaching, in particular in his presentation of the character of geography in relation to the whole field of knowledge.[79] Kaminsky's dissertation came to Hettner's attention in time to be mentioned in a footnote; that the geographic philosopher had come to the same concept as the philosophic geographer was recorded as welcome confirmation of the validity of the concept. In incorporating this essay in his volume of 1927, Hettner repeated this confirming footnote and added a paragraph in the text quoting from Kant.[80]

Three years after the publication of Hettner's volume of 1927, another doctoral dissertation, by Döring at Frankfurt, brought together for the first time the methodological statements that were scattered in Humboldt's various works.[81] While Döring compared these particularly with the views of Hettner and found them essentially similar, he did not look back to Kant's statement. Shortly thereafter, Humboldt's lectures of the winter of 1827–28 were published for the first time.

With all these materials at hand, it was possible for me in 1939, in presenting Hettner's statement of his concept, not only to add Kant's complete statement but also to draw on Humboldt for additional confirmation.[82] My conclusion that the three were in essential agreement in their view of the position and character of geography as a science has since been accepted by German writers.[83]

CONCLUSION

Kant was the first, so far as we know, to state the concept we have been considering. But his statement has had no *direct* influence in modern geographic thought—other than as a form of confirmation. It may have had an indirect influence, through partial and uncertain connections: in slight degree only through Ritter, possibly in greater degree through an effect on Humboldt's thinking and only thereby, and only possibly, on the thinking of Richthofen and Hettner. In each case, however, it is quite possible that there was in fact no connection. On the whole it appears probable that Humboldt's original statement, published in 1793, was independent of the concept which Kant had been presenting in lectures since as early as 1775 but which was not published until 1802.

While Hettner indicated that his concept was intrinsic in the development of the field and hence was at least in part present in the thinking of his colleagues, he was not aware of any particular source nor is it possible to trace his concept back to that of either Humboldt or Kant. Rather, if we may paraphrase his thinking, the concept existed in the historic development of the field from very early times; several or many students may have formulated it independently. Its present importance in the thinking of geographers of the world, however, is most largely due to the work of Hettner.

Regardless of by whom or when formulated, the concept was ignored when geographers considered their subject in terms of views of science transferred from other sci-

[78] "Neue Angriffe auf die heutige Geographie," *Geographische Zeitschrift*, Vol. 40 (1934), p. 382.

[79] Willy Kaminski, *Ueber Immanuel Kants Schriften zur physischen Geographie. Ein Beitrag zur Methodik der Erdkunde* (Dissertation; Königsberg, 1905), pp. 15, 39.

[80] *Die Geographie . . . , op. cit.*, p. 115.

[81] Lothar Döring, *Wesen und Aufgaben der Geographie bei Alexander von Humboldt*, Frankfürter Geographische Hefte (1931).

[82] *The Nature of Geography*, pp. 134–35.

[83] Ernst Plewe, "Vom Wesen und den Methoden der regionalen Geographie," *Studium Generale*, Vol. 5 (1952), pp. 411, 415; Hermann Lautensach, *Über die Begriffe Typus und Individuum in der geographischen Forschung*, Münchner Geographische Hefte (1953), p. 9.

ences, in particular those of the individual natural and social sciences. It met with receptive response only when geographers considered their subject in terms of its own intrinsic characteristics.

The intrinsic characteristics of geography are the product of man's effort to know and understand the combinations of phenomena as they exist in areal interrelation in his world. These characteristics are therefore independent of any particular concept of the subject; rather they form the empirical fact on which such a sound concept must be based. Acceptance of the concept is in no way essential to work in geography, but it is of value to those students who wish to understand the nature of the field in which they work in relation and comparison with that of other fields of knowledge.

In particular, geographers from early times have observed that work in their field differs from that in many other sciences in the following respects: (1) the fact that geography has no one particular category of objects or phenomena as its specific subject of study but studies a multitude of heterogeneous things as integrated in areas; (2) geography cannot be classified as either a natural science or a social science, nor simply as a bridge between the two groups, but rather must study combinations in which both kinds of phenomena are intimately intermixed; (3) study in geography requires the use of two markedly different methods of study: the systematic examination of certain categories of relationships over the world or any large part of it, in general or systematic geography; and the study of the totality of interrelated phenomena in particular areas, in special or regional geography; and (4) while geography like all other sciences is concerned with the development and application of generic concepts and general principles or scientific laws, it is like history in that it is also concerned in large degree with the knowledge and understanding of individual, unique cases.

As I hope to have shown in detail in a forthcoming study, the concept stated by Kant and Humboldt and more fully expounded by Hettner provides a reasonable explanation of these empirical facts about the field of geography. Hence it is appropriate to suggest, with Hettner, that this concept is not to be considered as the invention of any one man or of any small number of scholars, but rather as the more or less conscious recognition of countless geographers seeking a common framework of reference for their work.

THE ANTECEDENTS OF THE DOUBLE-PEN HOUSE TYPE[1]

MARTIN WRIGHT
Louisiana State University

SOUTHERN literature abounds with references to a distinctive type of folk house variously referred to as the "double-pen," "dog-trot house," "saddlebag house," and "three P's," to name a few. It is generally found in the hilly interiors of the Southern states and is recognized as belonging to the "old way of livin'." The numerous extant examples of the double-pen constructed of logs associate it with the period of pioneer settlement of the interior South.

This type of house may best be described by the folk terminology "two pens and a passage" (Fig. 1). Its basic form is two single pens (or rooms) separated by an open passage (the "dog-trot") and covered by a common roof. This simple structure is supported on foundation piers and generally has gable-end exterior chimneys; eaves which are deep enough to encompass the chimneys; and porches extending across the entire breadth of the house. Houses with more rooms have appendages added to the rear of the house but generally to only one of the two pens. The addition of more rooms usually results in a further rearward extension attached to the same pen, yielding in floor plan an L-shaped house with an open passage through the front center.

The basic form of this house was originally introduced into the pine forests of the Southeastern states as a log dwelling by westward-moving migrants whose cultural background was predominantly Scotch–Irish. As Hanna notes:

> The backwoodsmen were Americans by birth and parentage, and of mixed race; but the dominant strain in their blood was that of the Presbyterian Irish—the Scotch–Irish as they were often called. . . . Mingled with the descendants of many other races, they nevertheless formed the kernel of the distinctively and intensely American stock who were the pioneers of our people in their march westward, the vanguard of the army of fighting settlers, who with axe and rifle won their way from the Alleghenies to the Rio Grande and the Pacific.[2]

The foregoing is common knowledge regarding the double-pen and its spread in certain frontier areas. The antecedents of this building form, however, are obscured by the cultural maze of American colonization.

Most of the literature and all of the hill people consulted contributed to the prevailing confusion regarding the origin and/or evolution of this house type. The problem had presented itself earlier in a study of Southern log cabins.[3] Library research had indicated that the double-pen may have been a Southern invention since several Southern states claim the double-pen as an indigenous architectural form. The development of the type is attributed to constructional expediency by the following reasoning:

Single-pens were limited in size by the length of the logs which could be handled by the pioneers with their meager supply of tools. After construction, the corner-timbering of a log single-pen resisted the lengthwise addition of logs. Consequently, the erection of a large house necessitated the construction of two or more single-pens under one roof.

This method of enlarging a log dwelling and thus yielding a double-pen was used quite early in the history of American settlement. The earliest example is recorded in both Shurtleff[4] and the Pennsylvania guide.[5] The first pen of this house was built in 1654, the second pen in 1698, and in 1810 the intervening space was walled with stone. Shurtleff footnotes his text as follows: "This method of adding to a log cabin was common in all parts of the eastern United States in the nineteenth century." This house was a double-pen in 1698 and thus precedes the arrival of large numbers of Scotch–Irish in America by a full twenty years. On a chronological basis it would thus be impossible to credit the Southern hill culture, largely a Scotch–Irish development, with the original invention.

[1] The author wishes to thank the Office of Naval Research and the Ford Foundation for assisting field research in Louisiana and Scandinavia, respectively.

[2] Charles A. Hanna, *The Scotch–Irish; or the Scot in North Britain, North Ireland, and North America* (New York: G. P. Putnam's Sons, 1902), Vol. 1, p. 86.

[3] Martin Wright, "The Log Cabin in the South" (Unpublished M.A. thesis, Department of Geography and Anthropology, Louisiana State University, 1950), p. 42.

[4] Harold R. Shurtleff, *The Log Cabin Myth* (Cambridge: Harvard University Press, 1939), p. 172.

[5] *American Guide Series, Pennsylvania* (New York: Oxford University Press, 1940), pp. 62–63.

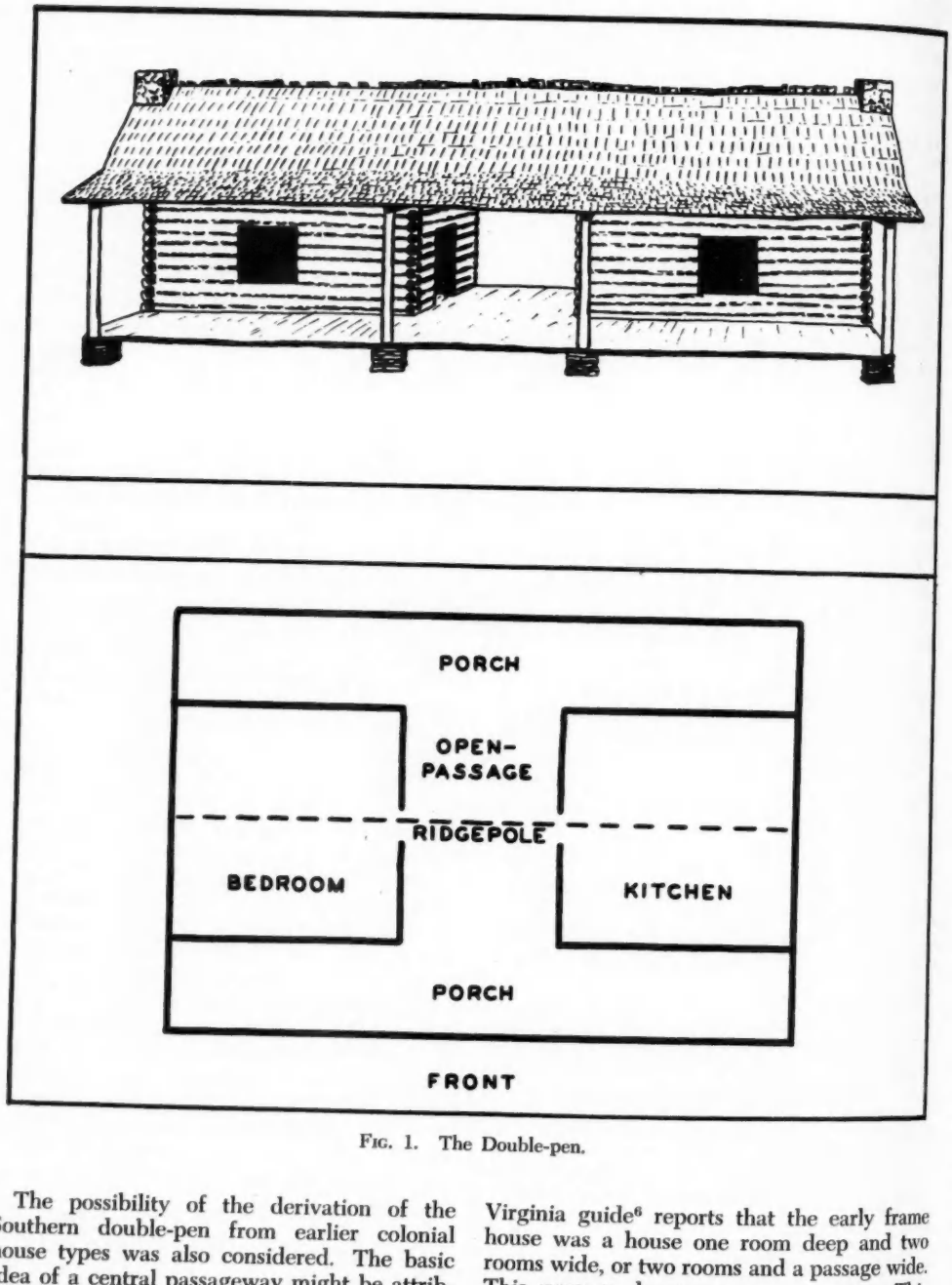


FIG. 1. The Double-pen.

The possibility of the derivation of the Southern double-pen from earlier colonial house types was also considered. The basic idea of a central passageway might be attributed to the influence of architectural types then found along the Atlantic seaboard—the Georgian house, for example. Further, the Virginia guide[6] reports that the early frame house was a house one room deep and two rooms wide, or two rooms and a passage wide. This passage, however, was not open. This simple description coincides quite closely with

[6] *Ibid.*, *Virginia* (New York, Oxford University Press, 1940), p. 175.

descriptions of frame double-pens found today in most of the South.

In similar vein, the floor plans of certain Irish and Welsh houses resemble closely the plan of the Southern double-pen. On the basis of floor plan alone, the possibility seems strong that the Southern double-pen is a modification of similar house forms known in Anglo-Saxon colonial America and in the western portions of the British Isles. However, one of the most diagnostic features of the Southern double-pen is its open passage. This aspect of a dwelling was foreign to the Anglo-Saxon cultures involved in the colonization of the Atlantic Coast of America.[7]

Other European culture groups were also represented in the east-coast colonies. One of the earliest permanent settlements of Europeans on the Atlantic seaboard was that of the Swedes who settled on the Delaware. Politically, this was a short-lived (17 years) venture. This colony was relatively small. Between 1638 and 1655 twelve expeditions were sent out from Sweden. The first landings were made at Christina Kil, on the west side of the Delaware River at the present site of Wilmington, where Fort Christina was founded. The eleventh expedition did not reach New Sweden, and the twelfth arrived in 1656—one year after the Dutch had taken over the colony. In 1653 the population of New Sweden was 70. In 1654 it numbered around 350, including some Dutch and a few Finns from Värmland.

Culturally, the colony had far-reaching effects. Extracts from Nelson are illustrative of the ecclesiastical influence of the Delaware colony and its effect on language and other traits:

> The English became successors to the Dutch as masters of the settlement, but the Swedes and Finns were allowed freely to develope their farming and their industries.
>
> As the Swedes were the most numerous and received their clergymen from Sweden, whereby service was obtained and the language was maintained, many of the Dutch and even other peoples to a great extent adopted Swedish language. . . . Dutch . . . English, Scotch, Irish and German families; they all used the Swedish language.
>
> [In] 1693 . . . 139 families and ninety-nine individuals desired to obtain sermon in Swedish. They were for the most part farmers. As for meat and drink they maintained Swedish customs.[8]

Swedish was still spoken on the banks of the Delaware 150 years after the arrival of the first Swedes and 137 years after New Sweden was taken over by the Dutch.

Of greater importance to the present study, however, is the notation by several writers of the construction of log buildings in the colony as well as an early resistance by two neighboring ethnic groups to the adoption of such construction methods. Jasper Danckhaerts, a Dutchman who toured the northern colonies in 1679, points out the relative superiority of the Swedish log cabins. He also points out the fact that the English resisted adopting the type despite the miserableness of their own dwellings.[9] The early Dutch also resisted accepting the type. Wertenbaker says, ". . . at tiny New Amstel on the lower Delaware the Dutch carpenters, scorning the log construction of the Swedish cabins nearby, went out into the woods axe in hand to trim beams for frame houses."[10]

The first ethnic group to adopt log construction extensively was the Scotch–Irish. These people began to reach the English colonies in strength in 1718 and by 1719 they were building log houses, a type unknown to them in their native country, but present in the Delaware settlement through which most of them entered America. It appears to have been this group who invented the term "log cabin" about 1750.[11]

Although the extent of Scotch–Irish borrowing of log techniques from the Swedes or from the Germans—who achieved a secondary introduction approximately fifty years later in interior Pennsylvania—was not determined, it seemed logical from the expressed influence of the Delaware colony (the major route of entry for the Scotch–Irish) that Scandinavia

[7] Numerous works including:
E. E. Evans, *Irish Heritage* (Dundalk: Dundalgan Press, 1943);
G. E. Fussell, *The English Rural Labourer, his home, furniture, clothing and food from Tudor to Victorian times* (London: the Batchworth Press, 1949);
Ake Campbell, "Notes on the Irish House, *Folkliv*, Vol. 1, nos. 2/3 (1937), pp. 207–34;
Sigurd Erixon, "Some Primitive Constructions and Types of Lay-out, With Their Relation to Rural European Building Practice, *ibid.*, pp. 124–55.

[8] Helge Nelson, *The Swedes and Swedish Settlement in North America*, Vol. 1 (Lund, 1943), p. 77.

[9] Shurtleff, *op. cit.*, p. 10.

[10] Thomas J. Wertenbaker, *The Founding of American Civilization: The Middle Colonies* (New York, 1938), pp. 46–7.

[11] Shurtleff, *op. cit.*, p. 26.

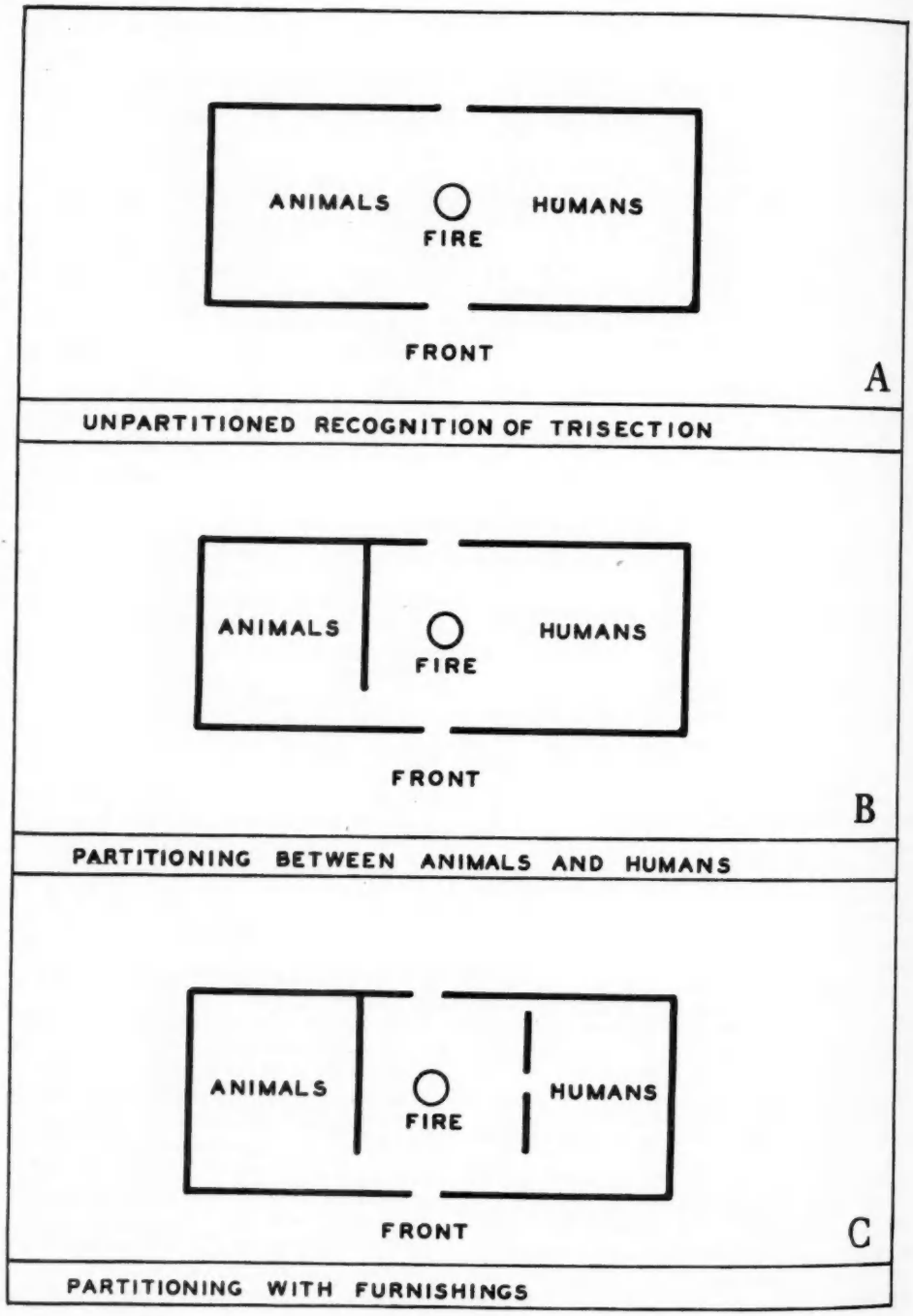


FIG. 2. Trisection in Primitive European Construction.

was the most likely European area in which to seek the antecedents of the double-pen. Strong support was given to this surmise by two reliable observers who had traveled in Finland and reported the presence there of barns built in the same design as the double-pen. Further support was obtained by personal conversation with Dr. Sigurd Erixon of the Nordiska Museet, Stockholm.[12] Dr. Erixon instantly recognized a simple sketch of the double-pen and termed it without hesitation a "pair-cottage."

Three features of the Southern double-pen proved to be of value in unraveling its origin from what at first appeared to be an anastomosing pattern of evolution, counter-evolution, and independent invention. These features were: (1) the open passage, (2) the gable-end chimney, and (3) the trisected floor plan. Rectangularity of floor plan was of no value. Rectangular plans were known in ancient days in the Mediterranean region, while oval and round plans were primitive forms in other areas of Europe. Rectangularity of plan was spread widely by the Roman legions, Christianity, and North Sea commerce,[13] so that by the time of the peopling of America the rectangular plan was used in all of those portions of Europe concerned with the colonization movement.

Primitive construction in Europe utilized quite widely the trisected floor plan. In the earliest primitive forms the trisection was not entirely apparent, although it was recognized that one particular end of the room was for animal occupancy, the opposite end was for humans, and the central portion was for the fireplace and the cooking (Fig. 2A). This recognition of a divided space soon led to partitions with various degrees of permanence. The most significant partitioning was that which separated human from animal, thus yielding a two-chambered floor plan with the smaller chamber occupied by the animals and the larger occupied by the humans and the fire (Fig. 2B). Though this two-chambered primitive form reached dominance in Europe, there was still a basic recognition of trisection, if only expressed by the arrangement of furniture[14] to separate the living quarters from the central fire chamber (Fig. 2C).[15]

A most significant point in primitive construction was the central location of the fireplace. Opposite doors were adjuncts of the central fireplace. The original purpose of the opposing doors was that of controlling the smoke from the fire. After losing this function due to the introduction of chimneys or stoves, the original plan of the doors was still retained.[16] The opposing doors of course allowed entry into the house from either the front or the back but, more significantly, they allowed passage through the house.

The movement of the fireplace to the gable end of the dwelling placed an increased emphasis upon the central chamber as a means of passage. In Ireland, a significant area for our purposes since it was the source region of the Scotch–Irish settlers of America, this trisected house plan, with emphasis upon the central chamber as a means of passage and the gable-end chimney, was a recognized cultural form at the time of America's colonization (Fig. 3A).[17] Retention of the fireplace in the central chamber resulted in a decreased emphasis upon the central chamber as a means of passage and finally a partitioning of the chamber to yield a kitchen in the rear portion, thus denying passage through the building. This resulting form is called by Erixon the "Frankish" type (Fig. 3B). It is widespread in Europe and extends into southern Scandinavia, particularly Denmark and the Swedish provinces of Bohusland and Scania.

In Sweden another important building practice was evolved. It was based on two important concepts. One was the very primitive method of house enlargement by the joining together of two similar basic units. Primitive dwellings of many sorts, including conical structures, were enlarged by building two similar units side by side, joining them together with a ridgepole, and then enclosing the enlarged area appropriately. I quote Erixon as follows:

> The Nordic system was forme[r]ly . . . quite different from this [the Frankish type]. The building design and the material [jointed timber] favoured a decided tendency to construct a number of small buildings arranged with regard to suitable corner

[12] Dr. Erixon granted several interviews in June, 1950, at Cooperstown, New York, while guest lecturer for the New York State Historical Society.

[13] Sigurd Erixon, "West European Connections and Culture Relations," *Folkliv*, Vol. 2, no. 3 (1938).

[14] Evans, *op. cit.*, p. 72.

[15] Campbell, *op. cit.*, pp. 211–12.

[16] Erixon, "Some Primitive Constructions and Types of Lay-out, . . . ," *op. cit.*, p. 142.

[17] *Idem*, and Campbell, *op. cit.*

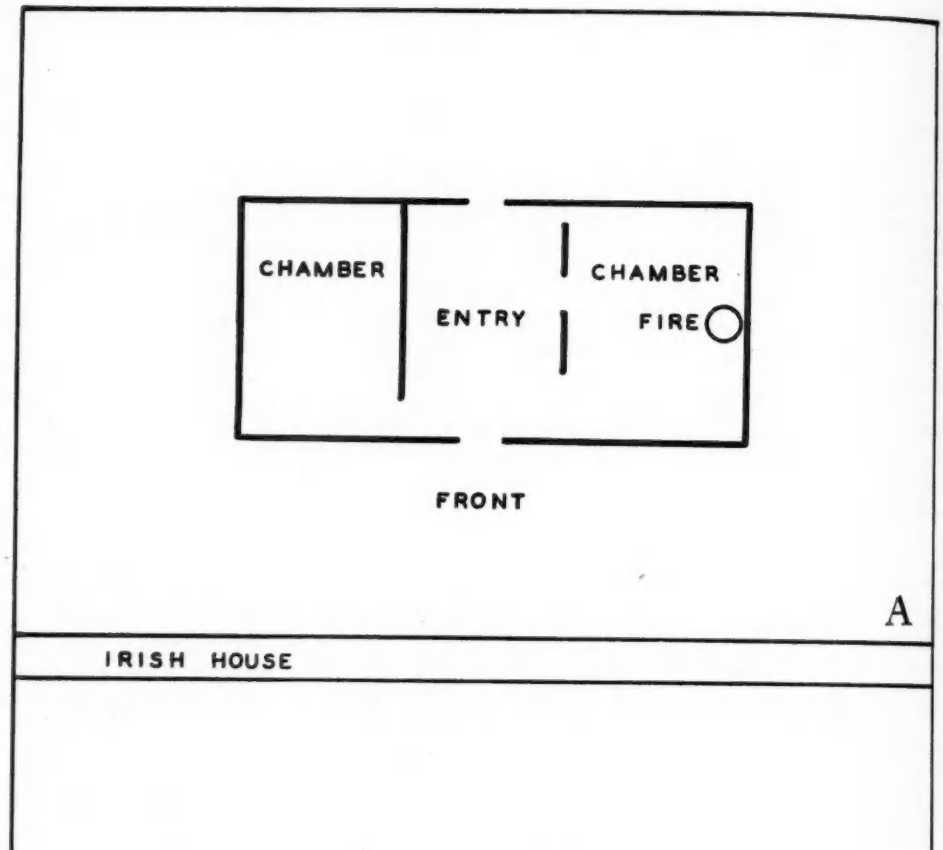


FIG. 3. Adaption to Fireplace Orientation.

distances for the timber. When larger buildings or several rooms were to be made, this was done by joining several houses. The Nordic pair-cottage illustrates this with especial clearness. . . . The two large side rooms correspond to houses still existing even independently in the country and comprise in one case a kitchen used as living room and in the other a store-room or best room.[18]

The second concept was Swedish recognition of the commonly accepted trisected floor plan. In a further comment by Erixon, attention should be called to the use of the word "imitation."

Everywhere in Scandinavia, Finland and other parts of the Baltic and neighbouring parts of Russia are to be found pair-cottages of this kind with entrance room having no fireplace in the middle. . . . Everything points to this house being a reflexion of the Continental, though the centre fireplace has been dispensed with. The imitation evidently took place by the joining up of two buildings already in existence, the fire-house and the store-house, both with gable doors, leaving an open passage between to be built in later on.[19]

Erixon brings out here another singular feature which the pair-cottage and the double-pen have in common. Doors in the interior gable ends of the individual pens, affording entry from the passages of double-pens, are not uncommon in Hill Louisiana (Fig. 4). In addition, Erixon says:

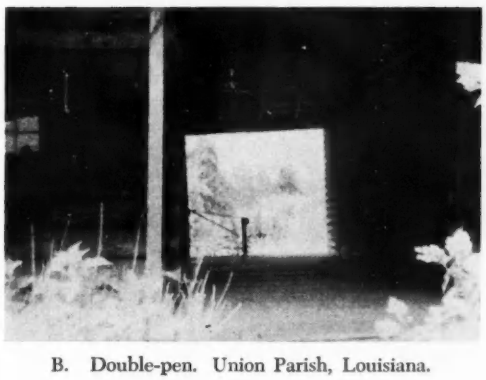
The result was a pair-cottage with an empty entrance room, open at both sides and unheated, which yet was gradually built and was partitioned to form a centre chamber. The fireplace was retained in deference to tradition in one of the side rooms. This introduction took place . . . evidently during the early part of the Middle Ages. Very often the one side room consisted on the Continent of a cowshed. This was replaced in Scandinavia, where people are less disposed than on the Continent to share their dwelling with livestock, by a store-room or a dwelling room.[20]

Although the development of the pair-cottage resulted in the presence of the fireplace in one of the side rooms of the Frankish type to the south, such fireplaces were not at the gable end of the house.

A detailed study of more recent architecture in the middle provinces of Scandinavia would be required to determine the degree of cultural induration of the idea of an open passage through a building. This idea is illustrated time after time in urban construction—even to the extent of routing the main thoroughfare through the city hall of Uppsala (Fig. 5A). A tantalizing similarity may be noted between the photograph of the Uppsala city hall and the photograph (Fig. 5B) of a folk house from eastern Norway. Further similarities may be noted between the photograph of a warehouse (Fig. 5C) in Stockholm and the photograph of a barn from Middle Sweden (Fig. 5D).

Thus, at the time of the European colonial effort in America, practically all of Europe was familiar with a trisected floor plan—one of the basic features of the double-pen. In addition the Anglo-Saxons, particularly the Scotch–Irish who figured so prominently in the settling of the South, were culturally adapted to the gable-end chimney—a second

A. Pair-cottage. Dalarna Province, Sweden.

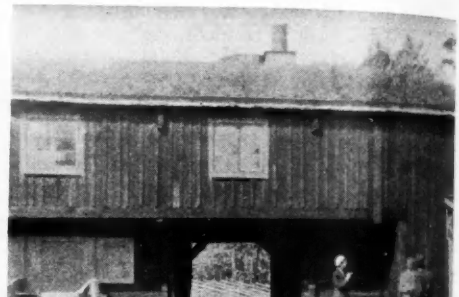


B. Double-pen. Union Parish, Louisiana.

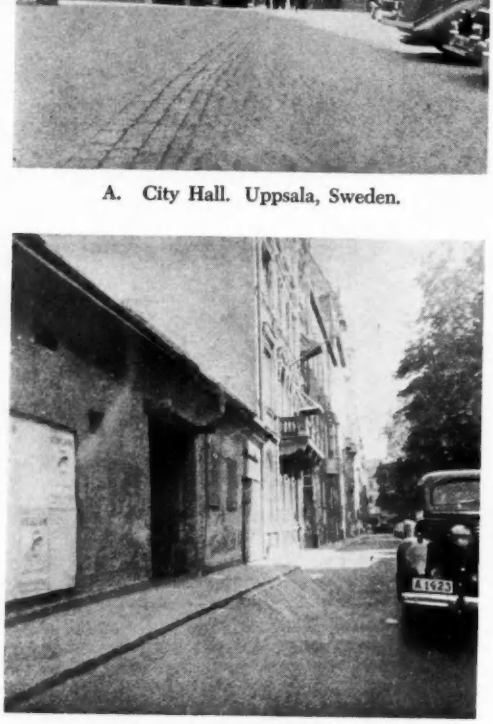
FIG. 4. Gable-end Doors.

[18] *Ibid.*, p. 153.
[19] *Loc. cit.*
[20] *Loc cit.*

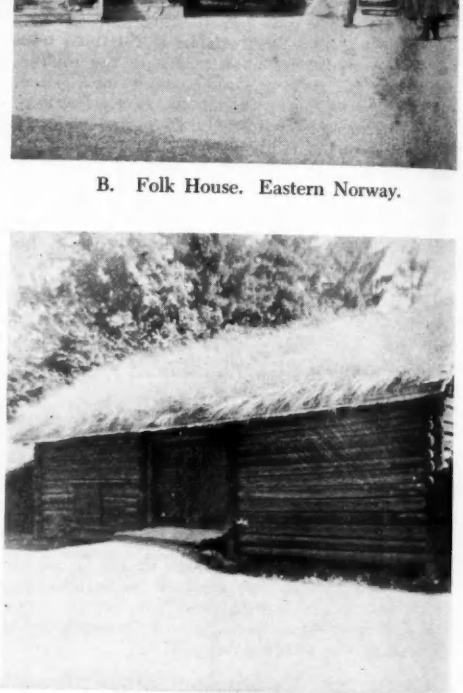
A. City Hall. Uppsala, Sweden.



B. Folk House. Eastern Norway.



C. Warehouse. Stockholm, Sweden.



D. Barn. Middle Sweden.

FIG. 5. Open Passages in Scandinavian Architecture.

diagnostic feature of the double-pen. Finally, the Swedes were familiar with the open passage as a feature of the basic house plan. The Scotch–Irish were not familiar with the open passage but they were acculturated to the notion of passage through a house. These were the significant cultural ideas prevalent at the time of the colonization of America.

It is an elementary assumption that the Swedish colony on the Delaware was just as much a pioneer settlement as were those of its English and Dutch neighbors, and a pioneer aspect of settlement features must have prevailed. By the time of the Swedish Delaware settlement, the Frankish house was established in the southern provinces of Sweden. To the north was the Swedish frontier region —the primitive or pioneer section—and in that area the pair-cottage had been an established feature of the cultural pattern since the early Middle Ages. There was (and still is) an association in the Swedish mind of "pair-

cottage" and "pioneer." In this country we have a similar association of "log cabin" and "pioneer." Many of the Delaware settlers came from the middle provinces of Sweden (e.g., the Finns from Värmland). It does not appear too naive to assume that at least one pair-cottage was built on the Delaware. It seems more probable that there were numerous examples of the type. Furthermore, the constructional progression of one unit, to the parallel unit, and finally the combination of the two, did occur as recorded in the case of the early double-pen mentioned above. It was the birthplace of John Morton near Chester, Pennsylvania, and corresponds by date and area with the Swedish culture.

Swedish influences still prevailed on the Delaware at the time of the earliest recorded date of double-pen construction within the present Southern states and long after the first large-scale arrivals of Scotch–Irish. The following quotation is extracted from the earlier library research:

> The earliest date of the construction of a double-pen in the south was 1776. It was built at Lawrenceburg, Kentucky, by one Coffman, a "Dutchman" from Pennsylvania. This report, in the Kentucky guide, implied that the double-pen was built as such upon Coffman's arrival at Lawrenceburg, indicating an earlier origin of the type.[21]

Of prime significance is the fact that there were definite early contacts between the Scotch–Irish and the Swedes. As noted earlier, the major route of entry of the Scotch–Irish into America was by way of the Delaware[22] and thus through the heart of the area influenced so strongly by the Swedish culture.[23]

In summary it should be reiterated that: (1) the open passage was a Swedish building feature; (2) the Dutch, Germans, and English were not only unfamiliar with the open passage but they also lacked the notion of passage through a house; and (3) the Scotch–Irish were accustomed to the "passage-through" notion as well as the gable-end chimney but were not familiar with the open passage. The Southern double-pen is thus a building form adopted by other ethnic groups from the Swedes on the Delaware and diffused most rapidly and widely by the Scotch–Irish to whom the open passage was not a too-greatly foreign concept and to whom may be credited the gable-end chimney modification.

This does not preclude the possibility of independent conception of the double-pen, since constructional expediency plus familiarity with the trisected plan could have led to the same results as achieved by the Swedes earlier. It is conceded that in isolated cases such independent invention may have occurred.

[21] Wright, *op. cit.*, p. 45.

[22] Hanna, *op. cit.*, Vol. 2, p. 37.

[23] Nelson, *op. cit.*, p. 77.

THE COLD SUMMER OF 1816

JOSEPH B. HOYT
New Haven State Teachers College

THE New England farmer of the early nineteenth century called the year 1816 "Eighteen hundred and froze to death." In this manner he indicated his attitude toward the extraordinary weather of that year. He saw it as a local tragedy or at most one confined to the northeastern part of the United States and was unaware that he was experiencing a phenomenon that was widespread in the northern hemisphere.[1] In the United States the year as a whole was cold, but the greatest hardship was caused by the exceptionally cold and dry summer months. Articles published on the year report monthly temperature only;[2] there are, however, a number of daily records available for New England.[3] Figure 1 indicates the climatic records used in this paper. An examination of these records shows that the major cause for the subnormal temperatures of the spring, summer, and early fall months was a series of cold waves that brought freezing temperatures to the entire northeastern part of the United States and Canada. (The average daily temperatures for four New England stations are shown in Figure 4, below.) The tragedy of these cold waves was their effect on crops and the consequent reduction of food supplies in the area. This study will be concerned with a description of the events of the summer in New England and the effects of the summer on food supplies as indicated by food price fluctuations and population movements, and a brief consideration of some theories as to the cause of the unusual weather.

FIG. 1. Sources and types of climatic data, 1816.

Although daily temperatures were somewhat lower than normal throughout the summer season, it was the totally unexpected late spring, summer, and early fall frosts that caused the greatest damage. An unprecedented series of cold waves, accompanied by high northwest or north-northwest winds, which at elevated and northern points brought snow, moved across the area every month from April through September. Frosts were recorded as far south as New Haven, Connecticut, on May 15 and 18, June 7, 8, and 11,

[1] Lorin Blodgett, *Climatology of the United States* (Philadelphia, 1857), pp. 147–8. "In England 1816 was almost as extreme as in the United States and the effect of the great degree of cold was quite as great. The observations made in the Royal Society of London show the spring and summer to have been 2.8° and 3.8° colder than average. Both were famine years, the last equally so on the continent in France and Germany though the Black Sea countries were unusually favored and productive."

[2] H. W. Dove, *Über die nicht periodischen Änderungen der Temperaturvertheilung auf der Oberfläche der Erde* (Berlin, 1840), Part I, p. 128; Lorin Blodgett, *op. cit.*; Willis I. Milham, "The Year 1816, The Cause of Abnormalities" (Annual Presidential Address read before the American Meteorological Society, Dec. 31, 1924), *Monthly Weather Review*, December 1924, pp. 563–70; Alfred J. Henry, "Abnormal Summers in the United States," *Monthly Weather Review*, August 1927, pp. 349–53.

[3] Hartford, Conn., meteorological record, published weekly in the *Connecticut Mirror*, January to December, 1816; New Haven, Conn., meteorological record kept by Professor Jeremiah Day, Yale University, manuscript copy at Yale; Brunswick, Maine, *Results of Meteorological Observations made at Brunswick, Maine, between 1807 and 1859*, by Parker Cleaveland, reduced and discussed by Charles A. Schott, in *Smithsonian Contributions to Knowledge, No. 204*, June 1867; Williamstown, Mass., meteorological record kept by Chester Dewey, *Memoirs American Academy*, Vol. 4, Part 2, pp. 387–92.

August 22, and September 27 and 28.[4] East Windsor, Connecticut, only 50 miles to the north, reported, in addition, frosts on May 30, June 10, July 9 and 10, and September 22.[5] Central and northern New England had even colder weather. Throughout the entire summer this area had only two periods of over three weeks without frost, June 11 to July 8 and July 10 to August 13.

WEATHER OF 1816

The year 1816 began with a relatively moderate winter. In Lancaster, New Hampshire, it was unusually mild with less than the normal amount of snow.[6] Southward in Connecticut a series of warm spells alternated with periods of cold. Spring came tardily. Robbins, in his diary, commented several times during April that "the vegetation does not seem to advance at all." This seems to have been due as much to drought, however, as to cold. The New Haven, Brunswick, and Williamstown weather records each reported the March and April precipitation as much below normal, 1½ to 2 inches in contrast to the usual 3 or 4 inches. Looking back on the early spring, the editors of the *Albany Advertiser* stated that they had "no recollection of so backward a season . . . the length and severity of drought checked progress of vegetation, grass withered."[7]

Lack of pasture made it necessary to continue stall feeding of cattle. Since hay was scarce, as was to be expected at the end of the winter season, corn was used, a necessity which reduced supplies for human beings.[8] This would not have been serious if the late spring months had been normal. In spite of the inauspicious conditions of April, farm work went on as usual. The Reverend Mr. Fogg of Kittery was out planting peas early in April and setting out apple trees later that month.[9] Mr. Robbins had his peas in and was clearing his asparagus bed by April 15 and cutting asparagus by the 30th. To the north, in Lancaster, New Hampshire, planting had been delayed. Adino Brackett did not begin his ploughing until April 29th; on the 30th he planted two and one-half bushels of rye.[10] Spring flowers bloomed late, the last of April in East Windsor; Robbins noted blossoms on peach trees and on daffodils the same day.[11]

The first part of May was cold and dry. The latter part was cold but the drought had broken. In Lancaster, Brackett noted: "The whole of this month has been so cold and wet that wheat could not be sown 'til late and then the ground could not be well prepared."[12] Fruit trees in the north did not blossom until late May, and in many sections the hard frosts of late May and early June caught the fruit trees, destroying leaves and blossoms and preventing the formation of fruit. Thomas noted this condition as far west and south as Cincinnati, Ohio.[13]

The lateness of the season meant that the frosts of mid-May did little damage since few planted crops were far enough along. Also, frosts were not uncommon in mid-May in northern New England. The unusual feature in the cold wave that appeared May 14th and lasted until the 18th was its southern extension to New Haven, and the accompanying snow which fell as far south as Chester, New Hampshire.[14] The frosts that came at the end of the month were more serious. Thomas, who was near Erie, Pennsylvania, on his westward trip, noted on May 29th: "Last night was cold. This morning . . . very frosty and ice covered the water ¼ inch thick. We had a brisk breeze from the northeast." On the 30th he wrote, "A severe frost attended this morning."[15] Rob-

[4] Jeremiah Day, meteorological record, *op. cit.* (hereafter cited as Day record). No pagination.

[5] *Diary of Thomas Robbins*, edited and annotated by Increase N. Tarbox, 2 vols. (Boston, 1886), Vol. 1, pp. 655–91.

[6] *Diary of Adino N. Brackett, 1807–17* (manuscript copy in the New Hampshire Historical Society, Concord, New Hampshire; no pagination), entry for April 22, 1816.

[7] June 4, 1816.

[8] *Daybook of Deacon Enoch Little 1816* (manuscript in the hands of Miss Ruth Sawyer, Warner, New Hampshire; personal communication).

[9] "Journal of William Fogg, 1813–21," *Old Eliot*, Vol. 3, 1899, pp. 108–12.

[10] Diary entry for April 30th.

[11] Diary entries for April 15th and 30th.

[12] Diary entry for May 24th.

[13] David Thomas, *Travels through the Western Country in the Summer of 1816* (Auburn, New York, 1819), pp. 72–3. (Hereafter cited as Thomas, *Travels.*)

[14] Day record. Normally the last spring frosts in New Haven are in April. Only twice in the last 84 years have there been May frosts, one on May 1st, the other May 3rd. Joseph M. Kirk, *Weather and Climate of Connecticut*, Connecticut State Geology and Natural History Bulletin No. 61 (Hartford, 1939), pp. 233–4.

[15] Thomas, *Travels*, pp. 34–5.

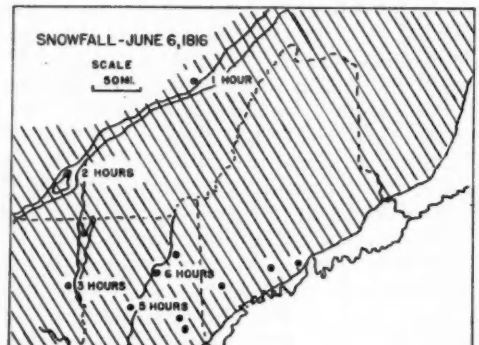


FIG. 2. Probable extent of snowfall, June 6, 1816.

bins, in East Windsor, expressed fear for the fruit trees. From the changes in barometric pressure and wind direction reported by Professor Day in New Haven it appears that the edge of the high pressure system extended that far south, although there was no frost. The wind shift, from southwest to northwest, came before sundown on the 28th, and the wind blew from the northwest quarter all the following day and into the 30th with a rising barometer.[16] Kingston, Rhode Island, reported similar winds on these days.[17] In Warren, Maine, the first fruit trees were set back greatly.[18] Corn was killed and had to be replanted in Norway, Maine.[19] Naturally there was much speculation as to the cause of these unusual weather conditions, and several observers noted and commented upon a large sunspot which had been visible to the naked eye during the first two weeks in May.[20]

Conditions seemed to improve the first few days of June. Temperatures were seasonally high and there were refreshing showers. The frosts of May 29th and 30th were forgotten, and men bent to repair the damage that had been done. Hope had revived too soon, however. Thomas, arriving in Pittsburgh in the afternoon of June 4th or 5th, described briefly the passage of a cold wave that was to bring disaster to New England: "On the day of our arrival in Pittsburgh we had several thundershowers from the west. The weather then became clear and for three days we had brisk gales from the northwest of unusual severity for summer . . . each night was attended by considerable frost."[21] When the cold wave reached the more northerly sections of New York and New England, it had become more intense. At Elizabethtown, New York, 130 miles due north of Albany, the front passed after dawn on June 6th:

> About 7:30 a.m. a storm of snow began which lasted three hours, succeeded by a strong westerly wind with some snow occasionally during the great part of the day. The wind prevented any white frost, yet the severity of the cold was such as to freeze the ground and destroy most of the garden vegetables.[22]

The wave continued eastward bringing bitter cold and snow to all the northeast. Evidently the front passed Montreal about 11:00 a.m.; a snowstorm began at that hour. Snow started to fall in Quebec early in the afternoon, indi-

[16] Day record. Professor Day includes a great deal of meteorological data in his daily reports including temperatures, wind direction, and barometric pressures taken at sunrise, 2 p.m., and sunset, as well as brief comments on the weather and amount of precipitation.

[17] Journal of Thomas R. Hazard, published as *Nailer Tom's Diary*, Introduction by Caroline Hazard (Boston, 1930). Weather data are scanty, consisting almost entirely of such words as "clear," "cloudy," "rain," and the direction of the wind. There is only one entry per day and no temperature report.

[18] Cyrus Eaton, *Annals of the Town of Warren, Maine* (Hallowell, Maine, 1851), p. 299.

[19] William B. Lapham, *Centennial History of Norway* (Portland, Maine, 1886), pp. 68–9.

[20] Fogg's journal, *op. cit.*, p. 109; *Farmer's Cabinet*, Amherst, New Hampshire, May 11, 1816, p. 3.

[21] Thomas, *Travels*, p. 55. The exact date of Thomas' arrival in Pittsburgh is not clear since he did not write the diary up every night. There is no entry between the 5th and the 9th. The quotation copied above was entered June 10th. The June 4th entry is dated "near Pittsburgh," but at the end of the entry are the two words, "arrive Pittsburgh." Since the front was traveling about 25 miles per hour between Elizabethtown and Montreal, a distance of 90 miles, it seems likely that it had been traveling at about the same speed between Pittsburgh and Elizabethtown. The airline distance between these two points is 420 miles. If the front had passed Pittsburgh June 5th in the early afternoon, this would allow it 16–17 hours to reach Elizabethtown by 7:30 a.m., June 6th, which was the time and date it was noted there.

[22] *Albany Advertiser*, June 22, 1816.

cating the passage of the front over that city.[23] In Lancaster, New Hampshire, it snowed not only on the 6th but on the 8th as well. Adino Brackett closed his diary account with the statement: "This is beyond anything of the kind I have ever known."[24] Temperatures were low enough to permit snow as far south as Amherst, New Hampshire, in the interior, and Warren, Maine, on the coast. Further south, snow fell only on the upland areas in the Berkshires. Figures 2 and 3 show the extent of the snowfall on the 6th and 8th.

The following four days, until June 11th, New England shivered, dug out winter clothing, rebuilt fires in fireplaces and stoves, and watched helplessly as the budding plants in the fields and gardens blackened under the icy northwest wind. No one was prepared; nothing could withstand it. Throughout the 7th the thermometer in Quebec stood constantly at the freezing point.[25] At Jackson, Maine, birds sought shelter in houses and barns.[26] Frosts were reported as far south as New Haven.[27] On the 8th it snowed again; amounts ranged from "a plentiful fall" in Portland, Maine, to six inches in Barnard, Vermont, and over a foot in Cabot, Vermont.[28] The high winds which accompanied the high pressure system moderated on the 9th and 10th. But the cold closed down. Frost was reported as far south as Worcester, Massachusetts, on the 9th.[29] The next morning, June 10th, an even more severe frost extended to East Windsor, Connecticut.[30] The severity of these frosts seemed to diminish from north to south, but everywhere vegetation suffered. Whatever corn had appeared was killed, as

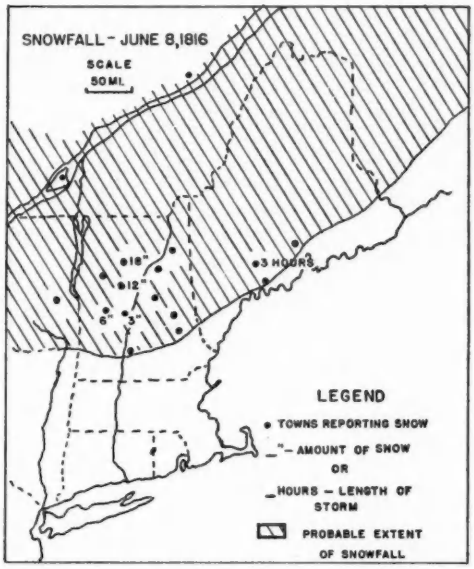


FIG. 3. Probable extent of snowfall, June 8, 1816.

were almost all the garden vegetables. "In northern towns many of the shorn sheep, although housed, perished and birds were frozen in their nests."[31] Thomas noted the effects of these frosts throughout Pennsylvania and Ohio until he approached the Ohio River at Cincinnati. He wrote:

> The late frosts have been very severe, we saw neither peaches nor apples 'til we approached the river [Ohio] and even here these fruits are scarce. Dead leaves in tufts are hanging on the papaw, and on most other trees. The first growth of this spring having been [*sic*] entirely destroyed. This remark will apply to much of the state where we travelled.[32]

The 11th saw the final frost of the cold snap, in many locations the most severe of the series, and then the temperature rose as the high pressure system passed, leaving the farmers to survey the damages and add up their losses. That these were considerable is evident in the gloomy note taken by most newspaper editors in reporting conditions. The comment of the *New Hampshire Sentinel* of Keene was typical: "Season very unpromising, we begin to despair of corn, hay will come extremely light."[33] Yet along with the cries of

[23] *Albany Advertiser*, June 19, 1816; *Connecticut Journal*, New Haven, Connecticut, July 6, 1816, p. 2, col. 5, report from Quebec.

[24] Entry for June 8.

[25] *Connecticut Journal*, New Haven, July 2, 1816.

[26] *Albany Advertiser*, June 28, 1816.

[27] Day record. Temperatures dropped steadily all day the 6th, the wind shifted from the west to the northwest between sunrise and 2:00 p.m., and the barometer rose steadily until noon of the 11th.

[28] W. M. Newton, *History of Barnard*, Vermont, 2 vols. (Burlington, Vermont, 1928), Vol. I, p. 285; *New Hampshire Patriot*, Concord, New Hampshire, July 10, 1816; *Windham* (Connecticut) *Herald*, July 4, 1816. Snow evidently blanketed New England north of the Massachusetts line.

[29] "The Diary of Isaiah Thomas, 1805–28," in *Transactions and Collections of the American Antiquarian Society* (1909), Vol. IX, pp. 299–337.

[30] Robbins diary, entry for June 10, 1816.

[31] *Albany Advertiser*, June 28, 1816.

[32] Thomas, *Travels*, pp. 105–6, entry dated June 28, 1816.

[33] July 10, 1816.

distress were sounds of confidence. The *Farmer's Cabinet*, another newspaper in the same town, stated that although the frosts had done great damage to the cornfields, "many have planted them anew."[34]

Four weeks of good growing weather followed except for the province of Quebec.[35] St. David had severe frosts on June 28th and 29th. But the farmer's trials were not yet over. The replanted vegetables seemed to grow with vigor; then, "when corn was being hoed down for the first time, there was a frost and cold sufficient to kill it down a second time."[36] On July 8th a new cold wave appeared, this one being reported from as far west as Chambersburg, Pennsylvania, and south to Richmond, Virginia.[37] Northern New England took the brunt of it, but even here it was not as severe as the June frost. None of the four stations for which we have daily records reported frost, although both Hartford and New Haven had temperatures in the forties. In East Windsor, Robbins noted that it was almost cold enough for a frost.[38] Even in northern New England some plantings escaped: "only those on low ground were killed or palsied."[39] In the uplands of central and northern New England, however, the new frost, combined with the drought that covered the area, "produced fears of a general famine."[40] From Kennebunk, Maine, reports stated that the hay crop was expected to be about one-third of the previous season's crop and that the frost had killed beans, cucumbers, and squash.[41] In Concord, New Hampshire, the estimate was a little more optimistic—one-half the usual crop of hay. A reduction in the hay crop meant that a corresponding reduction would have to be made in the animals kept through the winter.

A general shower on July 17th put heart back into the distressed population. The hardier grains, wheat and rye, were coming along well. Indeed, the cool spring was credited with having had a favorable effect on wheat. In St. David, Quebec, it was thought that the cool weather had destroyed several of the enemies of wheat, such as rust and the Hessian Fly.[42] By August 1st most editors had taken a new line and had begun to decry the tendency toward "famine fever." On August 6th the editor of the *New Hampshire Patriot* wrote:

> We are happy to announce to our friends at a distance that to the prospect of a general famine has succeeded one of uncommon plenty. The crops of hay are indeed light, rye is said to be better than for some years past, wheat and other early grains look well and are nearly ready for harvest, corn is more backward than usual although its growth has been luxuriant and rapid almost beyond example during the past three weeks so that there may be great crops even of the latter.[43]

These sentiments were echoed by other editors. The *Albany Advertiser* on August 5th commented on the promising crops and stated, "All apprehensions of a scarcity have subsided."[44] The editors did emphasize, however, the importance of continued warm weather.

It was not to be. In the middle of August, the 13th and 14th, frosts were reported in the interior of New York State and in northern New England.[45] Corn was again cut down; however, the frost was not felt south of Concord, New Hampshire, and southern New England had a week's respite. Another cold wave arrived August 20th, tumbling temperatures in Amherst, New Hampshire, some 30 degrees in a few hours. As the editor described it:

> A great and sudden change in the weather . . . took place on Tuesday. About noon a very violent storm of rain and wind went over. It came up very suddenly and was of short duration but it rained and blew tremendously accompanied by heavy thunder. After which the air became cold.[46]

North of Amherst, in Warren, New Hampshire, snow was seen on Mt. Moosilaukee.

---

[34] June 15, 1816.

[35] *Quebec Gazette*, July 25, 1816. (This information came from Jean-Charles Bonenfant, Library of the Legislature of the Province of Quebec.)

[36] Eaton, *op. cit.*, p. 299.

[37] *Farmer's Cabinet*, July 15 and 27, 1816.

[38] Diary entry for July 8, 1816.

[39] William Little, *History of Warren, New Hampshire* (Manchester, N. H., 1870), pp. 408–9; Charles C. Coffin, *History of Boscawen and Webster, New Hampshire* (Concord, N. H., 1878), pp. 184–5.

[40] *New Hampshire Patriot*, July 15, 1816. Three of our daily stations report drought in July: instead of the average 3 inches of rainfall, Williamstown had 2.1 inches, Brunswick 1.6 inches, and New Haven 1.2 inches.

[41] *Farmer's Cabinet*, July 15, 1816.

[42] *Quebec Gazette*, August 8, 1816; Thomas made a similar notation on the condition of wheat and rye in western Ohio. *Travels*, p. 102.

[43] August 6, 1816.

[44] It was comments such as these which misled Henry into thinking that the reports of the summer had been exaggerated. Cf. Henry, *op. cit.*

[45] *Albany Advertiser*, September 4, 1816; *Farmer's Cabinet*, September 7, 1816.

[46] *Farmer's Cabinet*, August 24, 1816.

Frosts were reported as far east as Portland, Maine, and as far south as East Windsor, Connecticut, during the next two days.[47] Travellers from Albany to Boston mentioned that "most of the corn between that city and Boston in low situations was destroyed. Most places have suffered severely from the drought."[48] This blow was followed by an even more severe frost on August 28th and 29th, which completed the work of destruction. In Keene it was said: "The severe frost of Wednesday [the 28th] has put an end to the hopes of many corn growers. Some whole fields have since been cut up for fodder."[49] In northern Maine the greater part of the corn was so treated in an attempt by farmers to salvage something for their efforts.

Again Nature seemed to relent, and at least the first few days of September were warm and sunny. The drought continued, however, and some localities in northern and central New England experienced frosts in the middle of the month. Hartford, Vermont, had a frost on September 10th; Concord, New Hampshire, on the 12th.[50] The drought broke in Connecticut with a heavy rain that hit New Haven on the 9th and continued almost steadily for eight days.[51] East Windsor and the rest of southern Connecticut had the same weather conditions, but farther north the drought continued.[52] Toward the end of September widespread frosts put an end to the growing season.

Of the harvest in general the comments from every side in New England were in agreement. Enoch Little of Boscawen said:

> Frosts killed almost all the corn in New England and not half of it fit to roast . . . on frosty ground the orchards were barren, but on warm land there was a moderate crop of apples . . . the prospects as to fodder are most alarming.[53]

Joseph Philbrick of Weare, New Hampshire, wrote that they had "a very small crop of hay, Indian corn the least known in the history of man, in consequence of the cold summer it could not ripen, but the crop of small grains was good."[54] From Montreal the report was that "Montreal district had a fair crop of wheat and a middling crop of other vegetables, but many parishes in Quebec must inevitably be in a state of famine before winter sets in."[55] In Concord, New Hampshire, an October newspaper account read: "Indian corn on which a large proportion of the poor depend is cut off. It is believed that through New England scarcely a tenth part of the usual crop of sound corn will be gathered."[56] Reverend William Fogg of Kittery, Maine, added another complaint in his September 14th diary entry: "It rains after the greatest drought known by our people. No prospect of crops. Crops cut short and a heavy load of taxes." Comments similar to the above appeared in every newspaper and in many diaries of the period. Adino Brackett summarized the personal reaction in his year-end entry: "This past summer and fall have been so cold and miserable that I have from despair kept no account of the weather. It could have been nothing but a repeatation [*sic*] of frost and drought."

Figure 4 shows the daily average temperatures at four stations in New England for the summer months of 1816. The cold waves described above may be clearly seen. The most noticeable is that of the period June 6th through the 11th. The only other one that shows up at all four stations is the cold wave of August 20th.

EFFECT OF THE SUMMER

A number of local historians commented upon the inflationary effect of the cold summer on the prices of agricultural products in the fall, winter, and spring of 1816–17.[57] But

[47] Robbins diary entry for August 22, 1816; Little, *History of Warren, New Hampshire*, p. 409.

[48] *New Hampshire Patriot*, September 17, 1816.

[49] *New Hampshire Sentinel*, August 31, 1816.

[50] W. H. Tucker, *The History of Hartford, Vermont* (Burlington, 1889), p. 13; *New Hampshire Patriot*, September 17, 1816.

[51] Day record. This was a northeast storm that dropped 5 inches of rain on New Haven in the eight-day period.

[52] Robbins diary. Precipitation records for Williamstown give August 1.7 inches and September 1.1 inches, as opposed to an average of 3 inches; Brunswick, Maine, had 2.1 inches in August and only 0.3 inches in September, in comparison with an August average of 4.4 and a September average of 3.0.

[53] Daybook, *op. cit.*

[54] Diary quoted by William Little, *History of Weare, New Hampshire* (Lowell, Mass., 1888), p. 371.

[55] *Albany Advertiser*, October 19, 1816.

[56] *New Hampshire Patriot*, October 22, 1816.

[57] Newton, *op. cit.*, p. 285: "Seed corn brought $5.00 a bushel in the spring of 1817."; Frederic P. Wells, *History of Newbury, Vermont* (St. Johnsbury, 1902), p. 263: "Corn sold for $2.50 a bushel, fall 1816"; E. L. Bogart, *Peacham, The Story of a Vermont Hill Town* (Montpelier, 1948): "Corn sold for $3.00 per bushel, 1817"; etc.

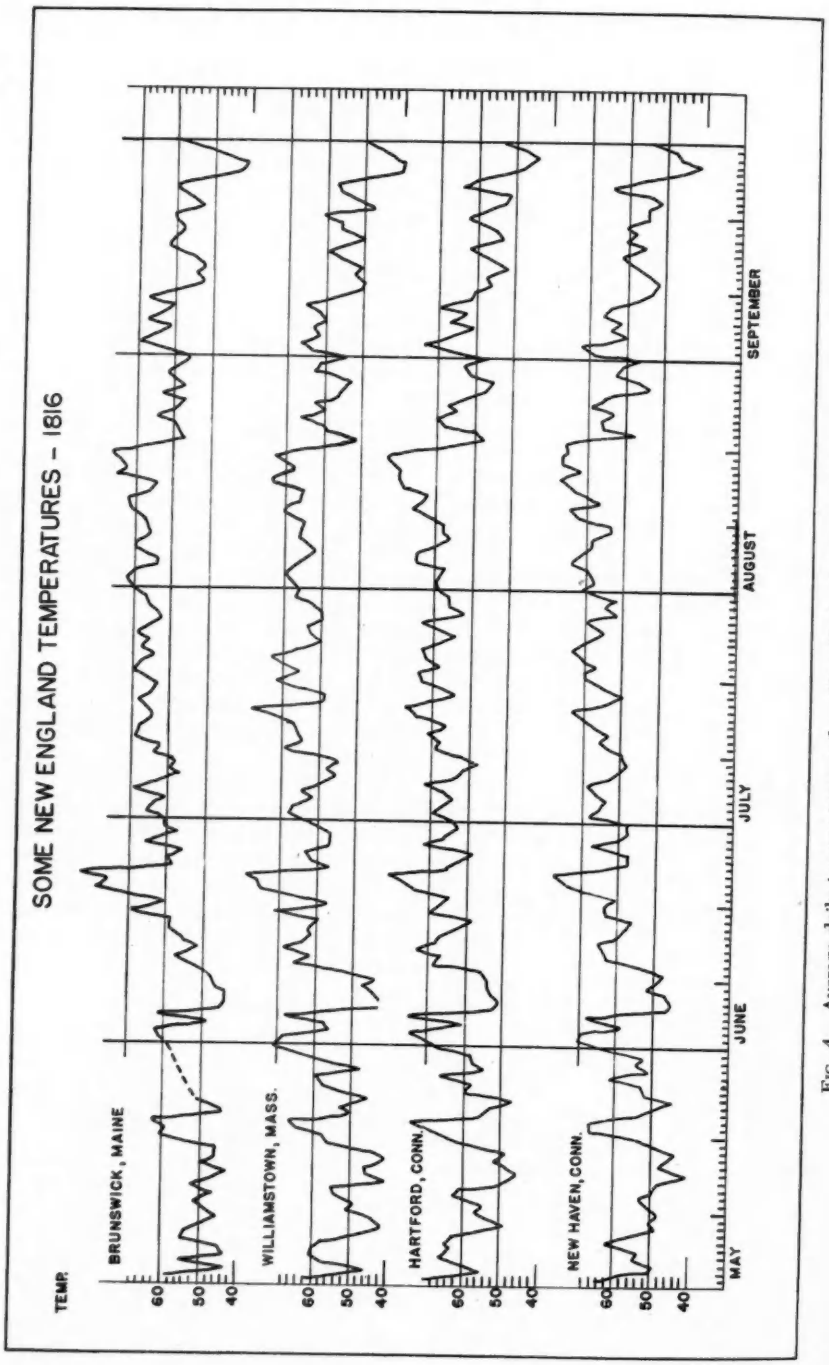


FIG. 4. Average daily temperatures at four New England stations, May–September, 1816.

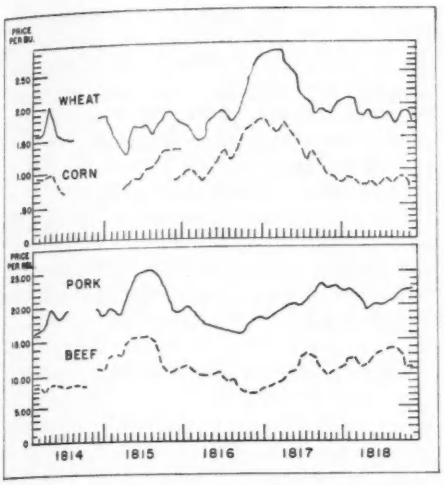


FIG. 5. Monthly average wholesale food prices, produce market, New York, 1814–18. The computation of average prices is explained in text footnotes 59 and 60.

their statements were valid only for the particular towns they were describing and did not necessarily reflect the situation in the northeast as a whole. It seemed that if the wholesale prices of either the New York or Boston produce markets could be charted, they would reveal what effect this summer had on the food supply of the region. Figure 5 shows the behavior of the prices of four major agricultural commodities, wheat, corn, beef, and pork, for the period January, 1814, to December, 1818.[58] Both grain and meat prices reflect the poor harvest of 1816 although in different ways.

Wheat clearly shows the effect of the poor harvest of 1816. With rare exceptions wheat prices fluctuated between $1.50 and $2.00 a bushel over the period January, 1814, to October, 1816. They returned to that price range after September, 1817. When the full extent of the disastrous summer of 1816 became evident in October of that year, the price of wheat began to rise. By January of 1817 it reached $2.75 a bushel and in April its peak of $2.87. Thereafter with the harvest of winter wheat in prospect, the price began to fall, reaching normal figures in the early fall of 1817.[59]

Corn behaved in almost exactly the same manner. With the exception of a few months at the end of 1815, corn prices usually ranged from $.75 to $1.12 a bushel. The June and July frosts of 1816 which threatened the fall harvest drove up the price of corn to $1.35 in August. Then the more favorable climatic situation of late July and early August, commented upon above, seemed to lift the specter of a poor crop momentarily, and prices sagged to $1.25. This same hesitation in rising prices in the late summer of 1816 can also be seen on the wheat chart. The continuing frosts of late August and September, however, changed the prospect of a bad harvest into reality and the market responded, driving prices to an average high of $1.78 by January and February of 1817. Like wheat, the price of corn fell to normal when the harvest of 1817 began to come on the market.[60]

The behavior of meat prices differed from that of the grain prices, but also reflected the cold summer of 1816. Normally the price of beef rose to a peak in August, then declined as farmers sent their livestock to market; this change was visible in 1815, 1817, and 1818. The prospect of a poor hay crop which became evident early in the summer of 1816 changed the usual pattern, and livestock began coming on the market late that spring and continued to arrive in large numbers throughout the summer. This wiped out the usual summer rise in prices. When the full picture of the disastrous harvest of 1816 became clear

[58] I have examined both the New York and Boston papers, but only the former gave a complete account of prices every week. Two papers were used: *New York Spectator* for 1814, 1815, and 1818; *New York Herald* for 1816 and 1817. Neither one gave a full five-year coverage for all items.

[59] The prices plotted are computed averages. Wheat was quoted each week with two prices, a high and a low, probably "asked" and "offered." There was fluctuation from week to week; for example, in July 1816, the week of the 6th, wheat was $1.81 to $1.87; on the 13th it rose to $1.87 to $1.94; on the 22nd it reached $2.00 to $2.03; and on the 27th it reached $2.06 to $2.09. This was computed as $1.81 to $2.09 and charted as $1.95.

[60] Corn prices were reported in the same way as wheat with the added complication of two grades being offered. These can be seen in the chart for December 1815 and January 1816 when "old corn" sold for $1.31 to $1.37, averaged as $1.34, while "new corn" was offered and sold for $0.87 to $0.94, averaged as $0.90. Also, beginning in August of 1817 and continuing, two grades, listed as yellow corn and corn (the former somewhat higher in price), are reported. Here the two are combined and the average is figured from the lowest to the highest price for the combination.

in September, the flow of cattle to market increased sharply. Many farmers, realizing the impossibility of keeping their stock through the winter, sent them off to market. Prices of beef in the fall and winter of 1816 dropped to a low of $7.75 a barrel. Breeding stock somehow must have been kept, since the prices in late 1817 and 1818 returned to the normal pattern, indicating no unusual scarcity following 1816.[61]

Pork prices followed almost precisely the same pattern as beef, the only differences being a lag of two months behind beef. Normally, pork prices seemed to reach their maximum in late September or October compared to the beef maximum in August. In 1816, farmers evidently began to ship their hogs earlier than their cattle. This drove the price of pork down to its low point a month earlier than beef. These minor variations, however, are not very significant. The important fact is that meat prices support the evidence of grain prices as to the poor harvest of 1816.

Other products listed by the New York commodity exchange showed the same trends. Rye followed wheat almost exactly, while butter, an animal product, indicated the reduced number of cattle in the summer and fall of 1816 by rising to a peak in the fall and winter of that year. These price fluctuations, important as they were in showing conditions in the region, were financial losses. The sufferings of the many subsistence farmers went largely unnoticed, although here and there in newspaper and diary accounts they were alluded to.[62] The over-all cost of the unusual summer could not be measured.

Although this paper has focussed upon 1816, that year was only one of a series of cold years, 1812 to 1818, and not even the coldest. The year 1817 bears that distinction.[63] However, in 1816 the summer temperatures were the lowest, and the effect on agriculture and the people the most disastrous. In an attempt to determine whether this series of cold years had any effect upon the well-known exodus from New England during the early decades of the 19th century the sources, diaries, newspapers, and local histories were examined for references to this migration. Only a few authors suggested this cold year as the cause. Typical of the comments were those from Barnard, Vermont, and Warren, New Hampshire, histories.[64]

The second possible source of information was the census records for the period. Accordingly the rates of population change were computed for each New England state for each decennial census from 1790 to 1860.

TABLE 1.—RATES OF POPULATION INCREASE, NEW ENGLAND STATES, 1790–1860

Area	Rates of increase, in percentages						
	1790 to 1800	1800 to 1810	1810 to 1820	1820 to 1830	1830 to 1840	1840 to 1850	1850 to 1860
United States	35	36	33	33	32	36	36
New England	34	19	13	18	14	22	15
Maine	57	51	30	34	26	16	8
New Hampshire	30	16	14	10	6	12	3
Vermont	82	40	8	19	4	8	0.3
Massachusetts	12	11	10	17	21	35	24
Connecticut	6	4	5	8	4	19	24
Rhode Island	0.4	11	8	17	11	18	36

These rates, given in Table 1, show that all the New England states but one, Connecticut, grew in population at a slower rate between 1810 and 1820 than in the previous decade. Also, every state except New Hampshire regained a more rapid rate of increase for the period 1820–30. Another interesting fact that appears is that the reduction in the rate of growth was greatest in the three northern states which felt the brunt of the cold: Maine, Vermont, and, to a lesser degree, New Hampshire.[65]

These three states had not even been fully occupied. In 1810 their population densities were 7, 22, and 23 per square mile, respec-

[61] Like wheat and corn, beef prices were quoted with a high and low weekly. The average has been computed on a monthly basis.

[62] Cf. footnotes 55 and 56.

[63] Average annual temperature at New Haven 1778–1935 was 50.0°. The years 1812–18 were considerably below this average: 1812—46.4°; 1813—48.5°; 1814—48.1°; 1815—46.8°; 1816—46.1°; 1817—46.0°; 1818—46.3°. Kirk, *op cit.*, pp. 91–3.

[64] Newton, *op. cit.*, p. 285: "When 1815, 1816, and 1817 all proved cold many migrated west." Little, *History of Warren, New Hampshire*, pp. 408–9: "So terrible was the year 1816 that the people grew disheartened and many sold out and went west and south."

[65] In southern New England three counties lost population between 1810 and 1820: Newport, Rhode Island; Litchfield, Connecticut, and Berkshire, Massachusetts. The last two counties are upland regions where the cold summer of 1816 would have been felt more severely.

tively. The three southern states had densities in 1810 of 52 for Connecticut, 57 for Massachusetts, and 63 for Rhode Island. Under the system of subsistence agriculture prevalent at that time, with an average-size family of 5–6 persons, this meant about 10 families per square mile or an average farm of 64 acres. If all the land had been good farm land and cleared, this would have been sufficient, but New England contained much land that was not arable, and the size of the average farm in New England at this period was 100–200 acres.[66] Allowance had to be made for the rocky, swampy, and sloping lands so frequently found in this area. Thus, southern New England had become overpopulated before 1810. The industrial revolution which was to provide work for the surplus farm population and markets for agricultural products in the new mill towns was still some years in the future. Between 1810 and 1820 the young man seeking his fortune migrated as he had been doing for several decades, the only difference being that after 1810 he went west rather than north.

The rate of population increase for the United States as a whole for the decade 1810–20 was 33.1 percent and this was almost entirely a natural increase, excess of births over deaths.[67] There is some reason to suspect a variation between birth rates in different parts of the country. Frontier regions tended to have young families, a larger percentage in the child-producing age, and fewer old people. Thus birth rates would tend to be higher in the more recently settled areas and lower in the older states along the eastern seaboard. The difference was not as great, however, as the difference in the rates of increase for the several states might indicate. The very rapid rates of increase for the decade 1790–1800, 82 percent in Vermont and 57 percent in Maine, were caused partially by higher-than-average birth rates—these were frontier regions at this period—but also by immigration from the southern New England states. On the other hand, the very low rates of increase for the same period in Rhode Island, 0.4 percent, and Connecticut, 5.5 percent, did not mean a reduction in births of this amount but rather emigration from the states. New England as a whole increased 13 percent for the decade 1810–20, or about 190,000 persons. Had it grown at the same rate as the rest of the country, it would have increased by 487,000. The difference between these two figures was the emigration from the area. Not that 297,000 persons migrated, but those who migrated, plus their progeny, totaled that figure.

Many factors entered into the migration from New England: exhaustion of the fertility of the rather thin forest soils, the attraction of cheaper and better soils in the western territories just opening up, overpopulation of some of the counties of southern New England, and others. The decision to move into a strange wilderness, to leave friends and family, to take the great gamble of a fresh start was one that thousands of families wrestled with. No single factor was the deciding one, but it seems logical that this series of cold years should bear some responsibility. After 1820 the migration continued but at a slower rate.

METEOROLOGICAL CAUSES

What happened in 1816 seems clear. A series of cold waves moved south from Canada and eastward across New England following the usual path of high pressure systems in North America. This is a common occurrence in the other three seasons, fall, winter, and spring. The problem is why they came in this summer season, and why they were so cold.

We do not as yet completely understand the cause of cold waves which originate, according to Henry, over the upper Mackenzie basin.[68] Tannehill advances the explanation of their origin one step further by suggesting that when the North Pacific is warmer than usual, air flows over the coastal ranges into the Alaska–Canada sink (Henry's upper Mackenzie basin) where the cold continental air mass is radiating its heat to the ground and, becoming colder and more dense, sinks to the ground, leaving room on top for the incoming Pacific maritime air. This process continues until pressure in Canada is much greater than in the United States, and cold air is forced

[66] Rudolph Dickinson, *Geographical and Statistical View of Massachusetts* (Greenfield, Massachusetts, 1813), p. 7.

[67] T. Lynn Smith, *Population Analysis* (New York, 1948), p. 308, suggests that the total immigration into the United States between 1783 and 1820 was less than 250,000, or 2.6 percent of the population in the latter year.

[68] Alfred J. Henry, "Whence Come Cold Waves," *Monthly Weather Review*, Vol. 56, No. 4 (April, 1928), pp. 143–5.

southward into the United States in the form of a cold wave.[69] But what causes these conditions is not yet clear.

Several authors have investigated a possible connection between sunspots and weather changes.[70] In general, their findings indicate that an increase in solar radiation follows increased sunspot activity. Insolation increases, at the equator especially, and the normal atmospheric circulation is intensified—more air flows north aloft, and high pressures are built up in the middle latitudes. Once formed, these high pressures tend to move in waves equatorward.[71]

One obstacle to using this general statement as an explanation for the cold summer of 1816 lies in the phrase "middle latitudes." Normal atmospheric circulation includes rising air at the equator and settling air in what become the high pressure systems centered roughly 30° north and south of the equator. If sunspot activity merely intensifies this primary circulation, then the high pressure systems referred to above start south of New England, and moving equatorward from there, would not affect that region. Tannehill gives a possible explanation.

The increased high pressures at 30° north would tend to increase the volume and velocity of both the northeast trades and the prevailing westerlies. The latter winds blowing from south to north in the northern hemisphere are warm winds. The prevailing westerlies in the Pacific are responsible for the North Pacific current, which is itself a warm current, especially the section flowing north along the British Columbian coast. The winds and current combined would warm the northeastern section of the Pacific Ocean. Tannehill suggests that the airflow from the ocean to the Alaska–Canada sink is increased when the ocean is warmer.[72] Thus we would have the necessary build-up of air in northwestern Canada to provide source material for the cold waves that assaulted New England in 1816.

In a later study, Clayton notes a relationship between marked outbreaks of sunspots and cold waves in the United States. Temperature changes, he believes, are influenced not so much by the absolute number of sunspots as by changes in the number. Measurements of temperatures in Boston made at the same time as the observation of sunspots show that the lowest temperatures reach that city about five days after the peak of sunspot activity.[73] Huntington notes a variation between the influence of sunspots on the rim of the sun and when they are in the center. In the latter position, they seem to reduce solar radiation rather than to increase it. Since sunspots appear on the right-hand rim and move across the sun, disappearing over the left-hand rim, there would tend to be fluctuations in solar radiation and correspondingly in the entire sequence of events, a phenomenon which would produce a series of cold waves. This is what happened.[74]

Sunspot activity was noted by the New England residents.[75] We have more accurate tabulations in the Wolfer sunspot numbers which show 1816 as a year of peak sunspot activity.[76] The Wolfer scale is an attempt to compute monthly and annual averages using the formula $N = k(10g + n)$ in which g is the number of groups of sunspots recognized on the sun at any one time counting isolated spots as groups; n is the number of spots which can be counted both within and without recognized groups; k is the multiplier not very different from unity depending on conditions of observation and the telescope used (if a large telescope, k is less than 1). It is clear that use of the Wolfer scale does not prove or disprove the Clayton thesis, and, lacking daily sunspot statistics, which are not available for 1816, it is impossible to point to sunspots as the only cause of the low temperatures of 1816. There does seem to be evidence, however, that sunspot activity may have been a partial cause.

Changes in solar radiation would, of course, affect temperatures in other parts of the world as well as in North America. Several observ-

[69] Ivan Ray Tannehill, *Drought, Its Effects and Causes* (Princeton, N. J., 1947), p. 91.

[70] These include, besides Tannehill: C. G. Abbot, numerous articles in *Smithsonian Miscellaneous Collections*; H. H. Clayton, *Solar Relations to Weather and Life* (Canton, Massachusetts, 1943); Ellsworth Huntington, *Earth and Sun* (New Haven, 1923); and others.

[71] Clayton, *op. cit.*, pp. 242, 286.

[72] *Op. cit.*, p. 91.

[73] H. H. Clayton, "Sunspot Changes and Weather Changes," *Smithsonian Miscellaneous Collections*, 1946, Vol. 104, No. 19, pp. 3–5.

[74] Huntington, *op. cit.*, p. 38.

[75] Cf. footnote 22.

[76] A. Wolfer, "Revision of Wolf's Sun Spot Relative Numbers," *Monthly Weather Review*, April 1902, pp. 171–6.

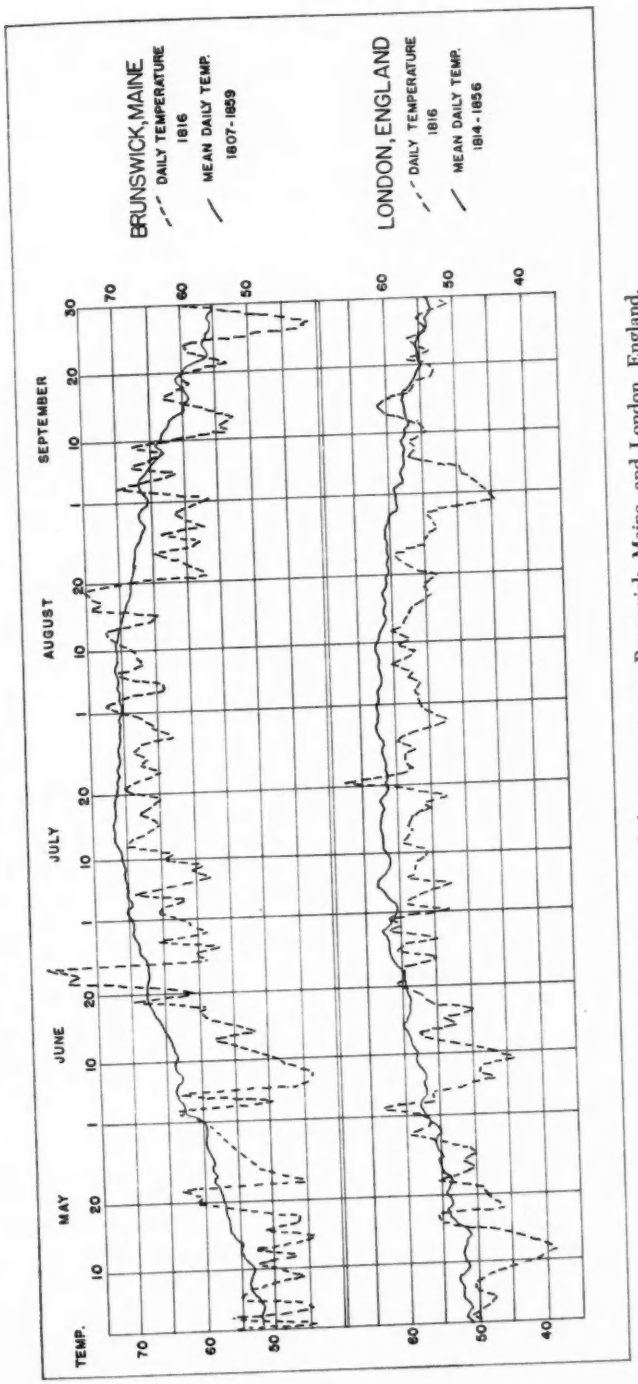


FIG. 6. Daily and mean daily temperatures—Brunswick, Maine, and London, England.

ers noted that the cold summer of 1816 was not confined to New England and adjacent areas.[77] Dove, who reported monthly averages for 34 European stations for that year, noted that with rare exceptions all had temperatures considerably below their means for each of the summer months of that year.[78] Since northwestern Europe normally has summer temperatures somewhat lower than the averages in North America, these temperature reductions would have been even more disastrous. He spoke of the year as *ein Jahr schrecklichen Miswachses* showing that crop failures were not confined to the American continent but extended to northwestern Europe as well.[79] Blodgett did suggest that the Black Sea countries were not affected by this cold season.[80]

To determine whether the cold summer in Europe was caused by the same or similar cold waves, the daily temperatures for London were graphed and compared to the mean temperatures for each day. These are presented in Figure 6, along with comparable data for Brunswick, Maine.[81] The resemblance between the two graphs is remarkable. In London, too, there was a series of cold waves that were largely responsible for the low summer temperatures. If one sets up the rather arbitrary definition that a cold wave exists if for two or more days the average temperatures are at least five degrees below the mean, Brunswick had thirteen such periods between May 1st and September 30th. During the same period London had nine, seven of which overlapped with seven of the Brunswick cold waves at least one day.

One other feature of these graphs is rather striking. At both stations the average daily temperatures for 1816 were frequently below the mean temperatures for the summer. London had only 16 days with temperatures above the mean for those days; in Brunswick there were 28 days that averaged above the mean. Seemingly there were other factors besides the cold waves which held temperatures down.

C. G. Abbot has presented another possible explanation for low temperatures: that of volcanic dust interfering with insolation.[82] It is well known that some of the greater volcanic explosions of the past have poured forth such volumes of material and thrown it so high in the air that some particles remain aloft for long periods of time. The haze formed by the volcanic dust of the eruption of Mt. Katmai on June 7, 1912, affected insolation as far away as Algeria. The reduction in insolation measured there increased steadily until mid-August and continued for some time thereafter. Abbot suggests that volcanic dust may reduce insolation as much as 10–20 percent, and may, if it continues long enough, reduce temperatures very markedly.[83] Again we lack exact data for this period, but it is worth noting that four of the greatest eruptions of history occurred in this period, Soufrière on St. Vincent Island in June, 1812; Mayon, Philippines, in 1814; Tomboro, Sumbawa, April 7–12, 1815; and Jan Mayen in 1818. All of these eruptions undoubtedly influenced temperatures throughout the world to a degree and for varying periods of time. The effects of the eruption of Tomboro were noted in the sunsets in Europe throughout the balance of the year 1815.[84]

SUMMARY

The summer of 1816 was cold not only in New England and adjacent areas but in northwestern Europe as well. As a result of the

[77] Blodgett, *op. cit.*, p. 147; W. H. Gardiner, "Record of the Weather from 1701 to 1885," *Special Papers of the Alabama Weather Service*, No. 1, no date; Dove, *op. cit.*, p. 128.

[78] Dove, *op. cit.*, Part I, pp. 78–9; Part III, p. 32; Part IV, pp. 78, 187. Of the 34 stations listed, seven were in Great Britain, seven in Germany, four in Russia (of these, two were in Finland), four in Austria–Hungary, three in France, three in Italy, two in Holland, one each in Norway, Sweden, Denmark, and Switzerland. The easternmost stations, St. Petersburg, Woro and Torneo in Finland, Archangel, and Stockholm each had temperatures above the mean in one or more of the critical summer months. Thus the cold may have been limited to countries of western Europe.

[79] *Ibid.*, Part I, p. 128.

[80] Blodgett, *op. cit.*, p. 147.

[81] Statistics for London, observations 1814–56, taken at the Royal Observatory, Greenwich, reported by James Glaisher in *Report of the Council, British Meteorological Society*, 7th Annual Meeting, 1857; for Brunswick, *Results of Meteorological Observations made at Brunswick, Maine, 1807–1859*, Parker Cleaveland, reduced and discussed by Charles A. Schott in *Smithsonian Contributions to Knowledge*, No. 204, June 1877.

[82] C. G. Abbot and F. E. Fowle, "Volcanoes and Climate, *Smithsonian Miscellaneous Collections*, Vol. 60, No. 29, pp. 7–10.

[83] *Ibid.*, p. 13.

[84] *Ibid.*, p. 4; also W. J. Humphreys, "Volcanic Dust and Other Factors in the Production of Climatic Changes," *Bulletin of the Mount Weather Observatory* (Washington 1913), p. 25.

cold, crop failures were widespread and suffering intense. In the United States, where populations were mobile, migrations to what were considered more favored regions increased. The cause of the cold temperatures seemed to be a combination of two factors: (1) varying solar radiation caused by sunspots and producing a series of cold waves that were unusually severe for the season and (2) a series of great volcanic eruptions which reduced insolation and temperatures generally throughout the period.

MAP SYMBOLS: EQUAL-APPEARING INTERVALS FOR PRINTED SCREENS[1]

ROBERT LEE WILLIAMS
Yale University

A common problem in map design is the one of selecting screens with which to show a scale of values. The screens should, as closely as possible, give a visual impression commensurate with the values they represent. A typical example would be a rainfall map with steps of 10, 20, 40, and 80 inches. How dark should each screen be to give the truest visual impression of these values relative to each other? The purpose of this study was to determine what values printed screens should have to give even-appearing visual steps from white to full color. Tone symbol tests were made with black, and, in order to see if colored screens behaved the same as black ones, for the colors red, orange, yellow, green, blue, and brown.[2] Tests were also given to see if the pattern of the screen would affect its visual evaluation, and tests were given to see what would be the effect of changes in the end points of the screen series.

The first problem in preparing these experiments was to draw screens that would print a known percent of the total area with ink. It was desirable to have screens with a wide range of tone. Two patterns were used, parallel lines and dots arranged in a triangular grid. These screens were drawn with the values shown in Table 1, the value representing the percent of the paper that is covered with the printing ink. The dot pattern was made as dark as possible without having the individual dots coalesce. Figures 1 and 2 show these screens.

Some preliminary testing showed that the intervals between the dark tones were too great. To correct this the screens for the triangular-spaced dots from 2 percent to 60 percent were photographically reversed giving additional screens with the following percent values: 40, 45, 50, 55, 60, 64, 68, 72, 75, 78, 80, 82, 84, 86, 88, 90, 92, 93, 94, 95, 96, 97, and 98. All of these screens were printed in

TABLE 1.—PERCENT VALUES OF SCREENS DRAWN FOR TONE SYMBOL TESTS

Percent of area in ink	Triangular spaced dots	Lines
2	X	
3	X	
4	X	
5	X	X
6	X	
7	X	
8	X	
10	X	X
12	X	
14	X	
15		X
16	X	
18	X	
20	X	X
22	X	
25	X	X
28	X	
30		X
32	X	
35		X
36	X	
40	X	X
45	X	X
50	X	X
55	X	X
60	X	X
65	X	X
70	X	X
75	X	X
80	X	X
85	X	X
90		X
95		X
100	X	X

[1] This material has been abstracted from a study I made with the aid of the office of Naval Research, Project No. NR 088-006 NONR 609(03), "Statistical Symbols for Maps: Their Design and Relative Values."

[2] In order to describe these colors in absolute terms, the tri-stimulus values of each color were determined by the Color Measurement Laboratory at the Massachusetts Institute of Technology. This specification of the colors used is important only to those wishing to duplicate them for comparison with other tests. Those wishing information on this method of color specification are referred to Arthur C. Hardy, *Handbook of Colorimetry* (Cambridge, Mass.: The Technology Press, 1936). The tri-stimulus values of the colors used for these tests are:

Red: $x = 0.380$, $y = 0.240$, $z = 0.073$
Orange: $x = 0.599$, $y = 0.372$, $z = 0.111$
Yellow: $x = 0.666$, $y = 0.714$, $z = 0.182$
Green: $x = 0.203$, $y = 0.323$, $z = 0.161$
Blue: $x = 0.207$, $y = 0.260$, $z = 0.518$
Brown: $x = 0.126$, $y = 0.1222$, $z = 0.091$

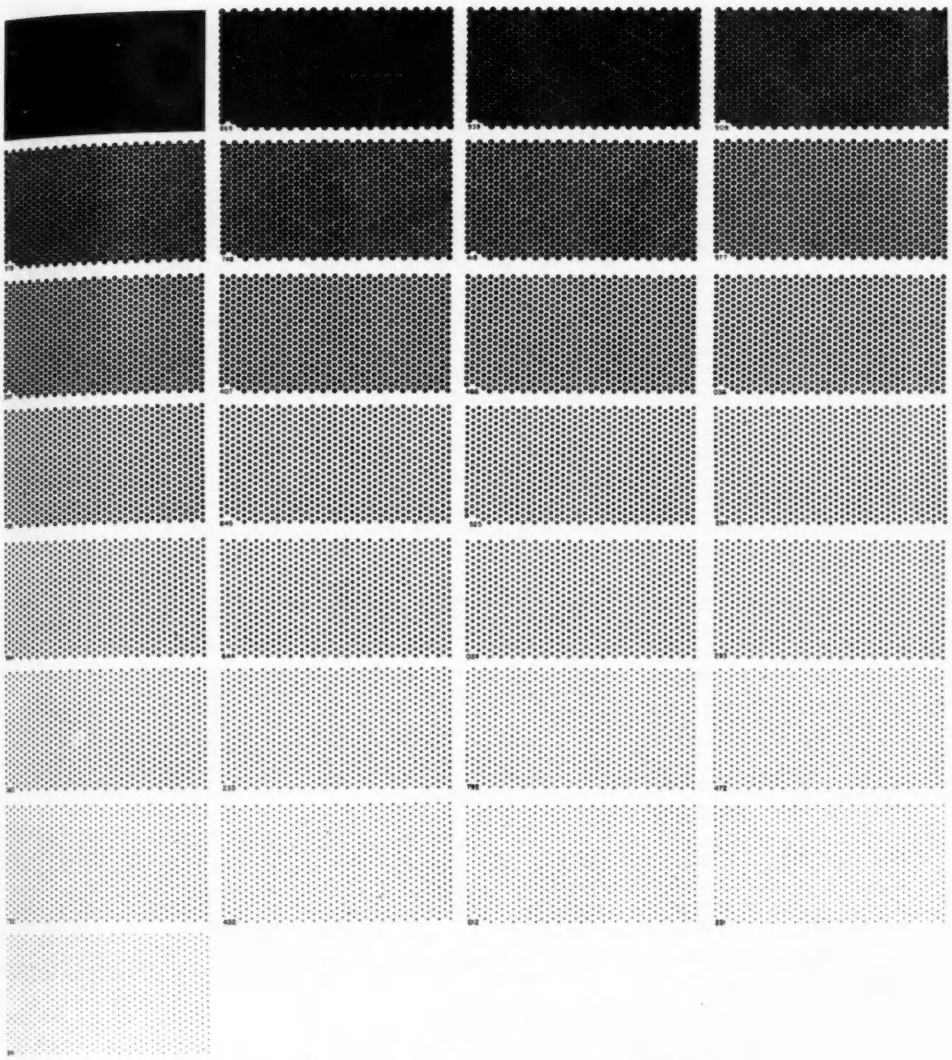


FIG. 1. Triangular patterned dot screens as drawn for the tests.

a haphazard arrangement on opaque white gummed paper and perforated in the manner of stamps so that they could be easily torn apart and pasted in place on the test sheets. They were printed in black and in colors.

BLACK SYMBOLS

Two groups of black tone symbol tests were given. The first contained seven tests covering the full grey spectrum from white to black. The second contained four tests dealing with segments of the full grey spectrum. Printed test sheets were used. The number of steps varied with the test sheets from three steps to eight steps, always with white at one end of the scale and full color at the other. One complete set of the tone stamps was supplied with each test sheet. From this stamp sheet selection was made, by trial, of the stamps that appeared to give even visual steps from white to full color and these were then pasted to the test sheet. Surplus stamps were

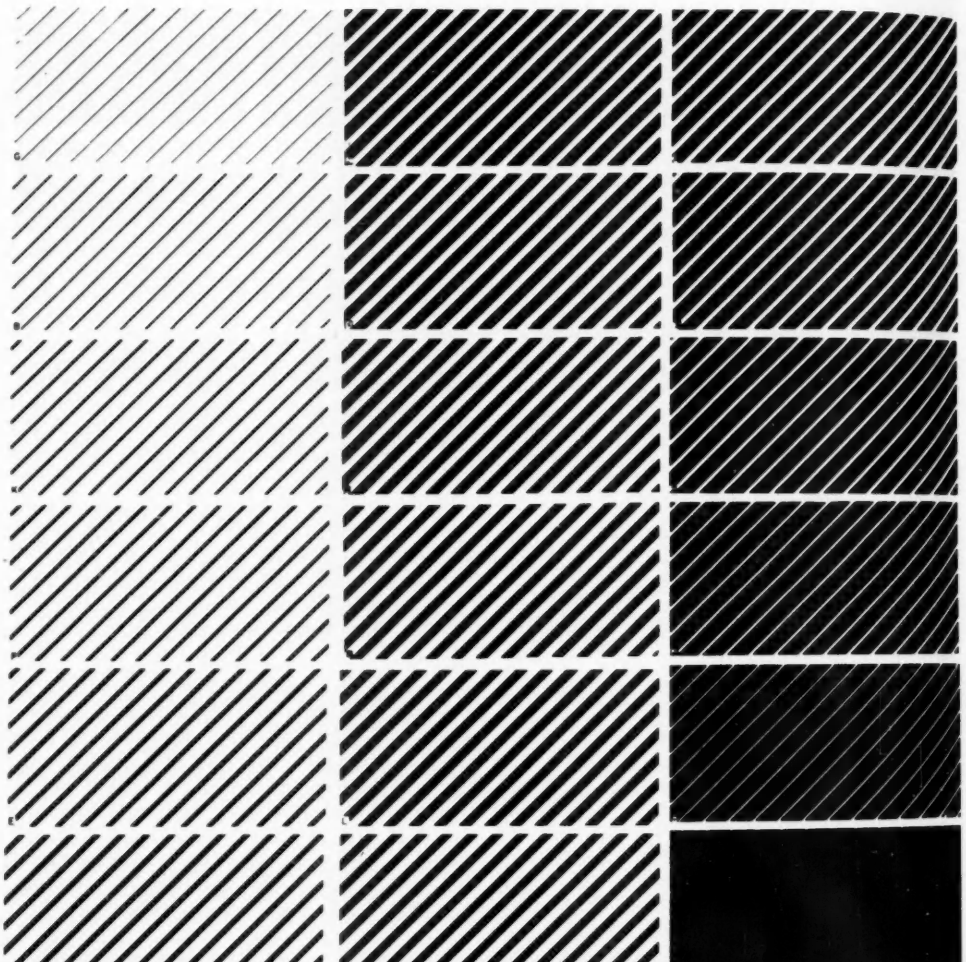


FIG. 2. Line screens as drawn for the tests.

discarded. The first group of tests was given to 50–60 sixth grade children, 23 cartography students at the University of Wisconsin, and 16 advanced design students at Yale. After the tests were given the answers were plotted, the median answer for each visual step was determined, and the medians for the sixth grade students, the University of Wisconsin students, and the Yale design students were averaged. These averages represent the percent of the area that should be covered with ink to give even visual steps from white (0%) to black (100%). The results are recorded in Table 2.

TABLE 2.—RESULTS OF BLACK TRIANGULAR SPACED DOT TONE SYMBOL TESTS

Visual scale interval	Percent of area that should be inked as determined by testing
Full color 100	100
85.7	92.3
83.3	88.7
80	85.3
75	80.0
71.4	76.3
66.7	67.7 and 69.3
60	56.0
57.1	54.0
50	47.0, 40.0 and 45.0
42.8	34.0
40	27.3
33.3	22.7 and 19.0
28.6	16.7
25	13.0
20	9.0
16.7	5.3
14.3	4.7
White 0	00.0

Figure 3 shows the curve obtained by plotting the results of the first group of tests on a logarithmic scale. This curve by not being a straight line shows that Fechner's law is not applicable to equal-appearing intervals of a grey scale (Fechner's law claims that sensation increases as the logarithm of the stimulus).[3] I have named this curve the curve of the grey spectrum.

[3] This contradicts the findings of Thurston. I think his error arose from testing only a small segment of the grey spectrum, 4 to 11 percent as closely as I can measure his samples, instead of the full spectrum, 0 to 100 percent. L. L. Thurston, "Fechner's Law and the Method of Equal Appearing Intervals," *Journal of Experimental Psychology*, Vol. 12 (1929), pp. 214–24. See also F. M. Urban, "The Weber–Fechner Law and Mental Measurement," *Journal of Experimental Psychology*, Vol. 16 (1933), pp. 220–30.

To see if the pattern of the tone screen would affect their selection, a test was given using line tone symbols. In this test the tones were composed of lines instead of dots and the steps between the tones were of even 5 percent intervals instead of the close intervals of the dot tone symbols. In spite of these differences from the dot screens, the result of this test corresponded closely with the results of the tests using the dot screens. The test asked for the selection of five steps between white and full color. Figure 4 shows these values plotted against the curve of the grey spectrum. It thus appears that the pattern of the tone symbol does not materially affect the selection of equal-appearing steps between white and a full color.

Four additional tests using black screens were given, using two screens as end points of the tone scale instead of white and black. This was done to determine whether a change in the tone of the end points would change the curve of the grey spectrum. These tests covered the portions of the grey spectrum as listed below:

Test 1: 3% black–three steps–45% black
Test 2: 45% black–three steps–85% black
Test 3: 3% black–three steps–85% black
Test 4: 18% black–three steps–70% black

Except for the difference in the end points, these tests were identical to the ones testing the full grey spectrum from white to black. The dot screens were used. The tests were given to 18 advanced design students at Yale and the medians of their selections were taken to be the best expression of the group. Figure 5 shows the results of these tests plotted against the curve of the grey spectrum. The close correspondence of the plotted points on these four graphs to the curve make it evident that whether the end points of a grey

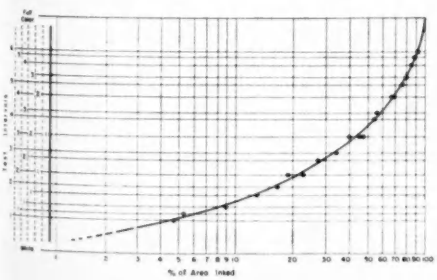


FIG. 3. Curve of results of black dot symbol tests plotted on logarithmic scale.

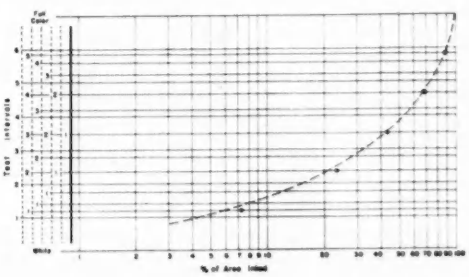
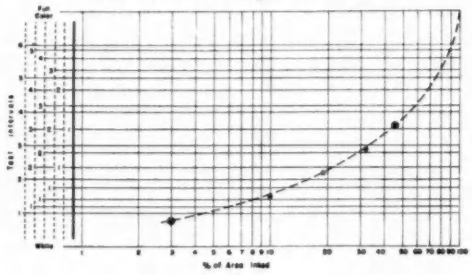
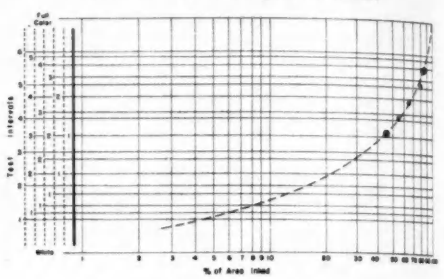


FIG. 4. Values obtained from line tone test plotted against curve of grey spectrum.

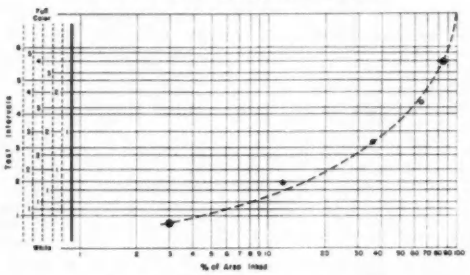
TEST 1. 3 to 45 percent black.



TEST 2. 45 to 85 percent black.



TEST 3. 3 to 85 percent black.



TEST 4. 18 to 70 percent black.

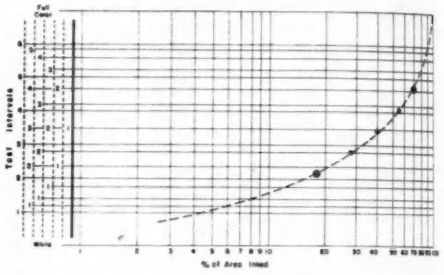


FIG. 5. Results from the tests to determine the effect of change in the end points of the tone scale. Values are plotted against curve of the grey spectrum.

scale are white, black, or tones of grey, the curve of the total grey spectrum as obtained from the dot tone tests is valid.

Figure 6 is an enlarged drawing of this curve of the grey spectrum from which equal-appearing tone steps may be measured. To do this, decide which end points are desirable for the tone scale wanted and plot them on the curve according to their percent of inked area to the total area. Between these two end points draw as many evenly spaced lines parallel to the base of the graph as you desire to

TABLE 3.—RESULTS OBTAINED FROM THE COLORED TONE SYMBOLS TESTS

Visual scale interval	Percent of area to be covered by ink						
	Black	Red	Orange	Yellow	Green	Blue	Brown
Black 100							
85.7	92.3						93.3
80	85.3	86	90.3	85	83	80.7	82.3
71.4	76.3						77.7
66.7	68.5			73			
60	56.0	53.3	56	64	53.3	51.7	54
57.1	54.0						55
42.8	34.0						37.7
40	27.3	28	29	40	28.3	25.3	27.3
33.3	20.8			32			
28.6	16.7						18
20	9.0	9.3	10.3	17	9.7	9	8.7
14.3	4.7						6.3
White 0							

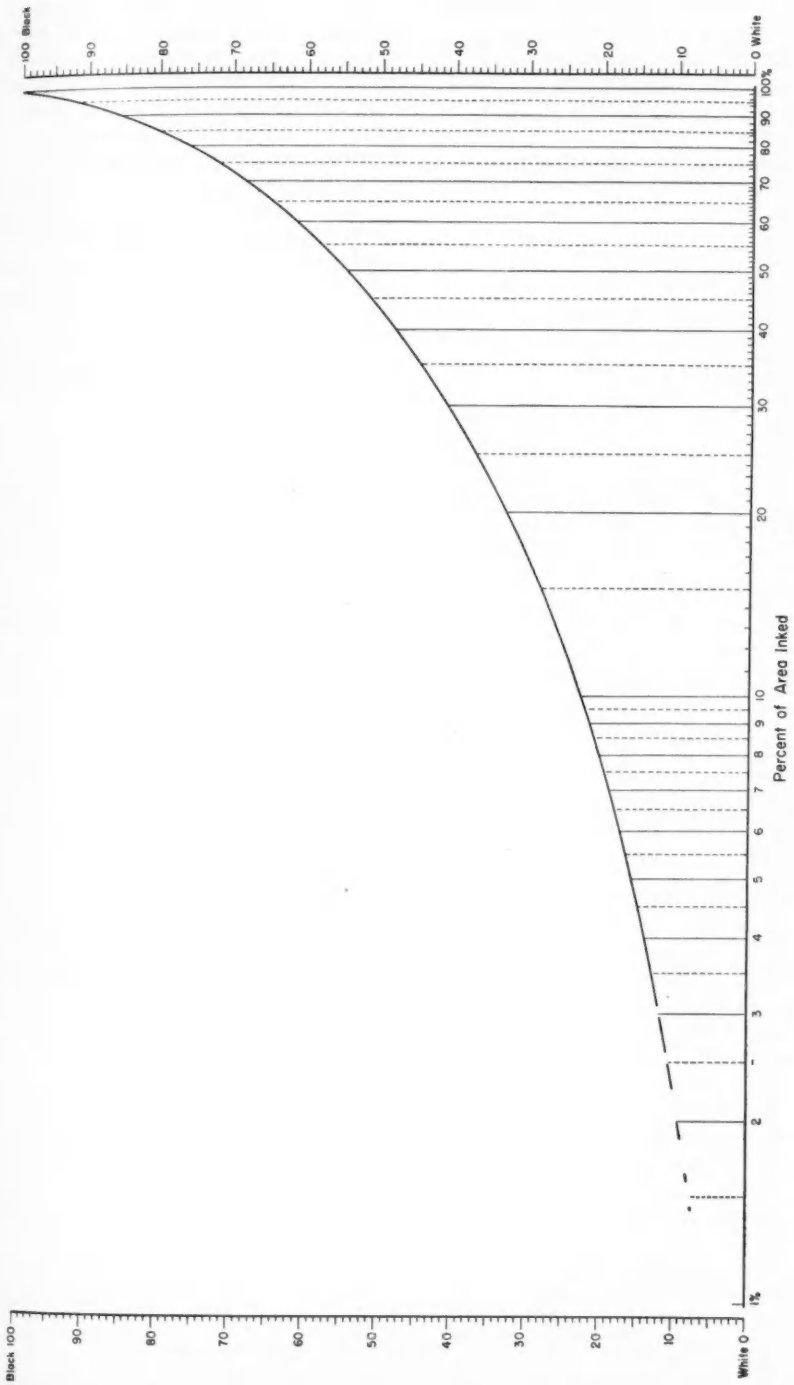


FIG. 6. Curve of the grey spectrum.

have tones between the end points of the grey scale. Where these evenly spaced lines intersect the curve, read the percent value of the screens that will give even appearing intervals to the grey scale. For example, if it is desired to have nine tones between white and black, the above procedure will give the following percentage value: 2.5, 8, 17, 28.5, 43, 58, 73, 85, and 94.

COLORED SYMBOLS

Eight tests were given using colored dot pattern screens:

Red: white–four steps–full color
Orange: white–four steps–full color
Yellow: white–two steps–full color
Yellow: white–four steps–full color
Green: white–four steps–full color
Blue: white–four steps–full color
Brown: white–four steps–full color
Brown: white–six steps–full color

The test sheets were identical with those used for the black symbol tests and were given to the same groups and at the same time as the tests on the black tone symbols. The results of the tests were graphed and the calculations made in the same way as for the black symbol tests. These results are shown in Table 3. Figure 7 shows these same results in graph form. On these graphs the curve of the grey

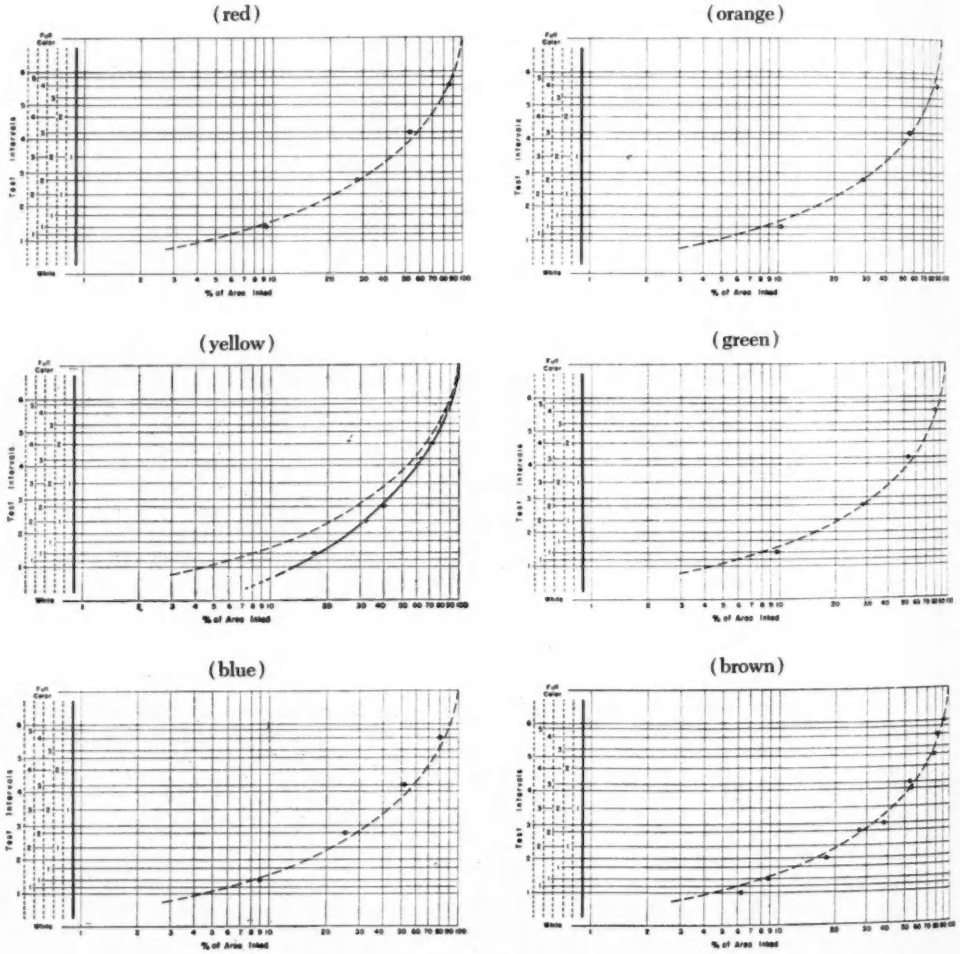


FIG. 7. Results of the colored tone symbol tests plotted against curves of the grey spectrum.

spectrum is shown by the dashed line and the values obtained from the colored tone symbol tests are shown as dots. It is immediately seen that for all except yellow, the curve for each colored tone symbol corresponds closely to the curve of the grey spectrum. Thus, to obtain even visual steps for any of these colors, except yellow, the grey spectrum curve may be used. For yellow the solid curve may be used to obtain even-appearing visual steps, but it would probably be better to avoid using such a light color for tone symbols.

CONCLUSIONS

1. Fechner's law is not applicable to equal-appearing intervals of a grey scale. The plot of the grey spectrum on a logarithmic scale is not a straight line, but rather a curve increasing in steepness toward the blacks.
2. The pattern of a tone symbol does not materially affect the selection of equal-appearing steps between white and a full color.
3. Whether the end points of a grey scale are white, black, or tones of grey, even-appearing steps between these end points will fall on the curve of the grey spectrum.
4. Except for light colors, particularly yellow, even-appearing steps of a colored tone symbol may be determined from the curve of the grey spectrum.

DECISION MAKING IN REGIONAL CONSTRUCTION[1]

LEONARD ZOBLER
Barnard College, Columbia University

SELECTING ACCEPTABLE BOUNDARIES

REGIONS are areal systems based on levels of similarities and differences in spatially distributed traits. Regional boundary lines are conceptual areal frames fitted to the earth's surface within which landscape features are sorted. It is the nonhomogeneous distribution of these features which identifies the peculiar properties of respective regions. Each spatial unit is internally homogeneous but differs from other members of the system.

The construction of a regional system entails the location of boundary lines. Geographic boundaries are located commonly in three ways:

1. By means of traverses, points are set at places where differences of acceptable magnitude occur in observed landscape features. The points then are connected and the resultant areal divisions will have different counts of the observed items. To avoid interpolation an infinite number of traverses must be run for each trait.
2. Unit areas, as counties or states, for which aggregate data are available, are compared with each other and then grouped into larger units. The process of combining rests upon a statistically calculated level of difference based on one or several indexes.
3. Areal units may be established by employing any continuously distributed landscape feature which the investigator believes is related to other traits in which he is interested. The features may be climatic, physiographic, political, etc., while the investigator may be interested also in the distribution of ethnic groups, land use, type of industry, etc. After the spatial pattern has been set up, data may be collected from the unit areas and examined to determine if acceptable levels of difference are found.

Whichever method is used, the investigator is faced with the necessity of making decisions among a set of proposed boundary lines. The basis for these decisions rests upon an evaluation of the sorting influence which the set of lines has upon the distribution of traits among the unit areas. Acceptable levels of difference should be defined objectively. There are several aspects of this problem. Is a difference of 2, 10, 20, 50 counts or percent adequate to state that the unit areas differ? If the level of difference is set after the data have been examined, the investigator may have been guilty of bias, unless certain procedures, as sequential sampling, are followed. Complete observations of the traits must be made, for if any are omitted there is no way of knowing how the results might have been influenced. Complete coverage, however, might not be feasible.

In procedures "1" and "2" above the distribution of landscape features determines their own "natural" bounds by speaking for themselves.[2] There is a sort of circuitous reasoning in these completely empirical methods which guarantees that differences will be found. These operations of region building are subjective, for the requisite differences can be set at any level. As a result such procedures make only a limited contribution to an understanding of the dynamics whereby areal differentiation occurs.

The use of statistical methods for the construction of regions is more explicit and more efficient for decision making than the methods commonly followed. Procedure "3" above is adaptable to testing operations. According to its methodology the precise boundaries of the areal units are established independently of the observations and before the test samples are taken. This independence enables the investigator to erect a hypothesis of relationship between the regional construct and the spatial distribution of selected landscape fea-

[1] This report is an elaboration and combination of papers presented at the annual meetings of the Association of American Geographers, April 2–5, 1956, Montreal, Canada, and April 1–4, 1957, Cincinnati, Ohio. The writer wishes to acknowledge helpful comments by Dr. W. L. Garrison of the University of Washington.

[2] W. S. Peters, "A Method of Deriving Geographic Patterns of Associated Demographic Characteristics within Urban Areas," *Social Forces*, Vol. 35 (1956), p. 62.

tures. In effect, the investigator is saying that he has drawn his boundary lines so that they have divided the traits into different areally distributed populations. He is postulating also that the basis of the boundary lines is related directly or indirectly to the traits. This is the step which demands geographic insight and contributes to an understanding of regional dynamics. The crucial question is, do the boundaries sort the traits into different populations?

Quantifiable sample data are collected from each of the unit areas and treated statistically to evaluate the hypothesis of relationship. As a result of the test it is possible to decide if the samples drawn from the unit areas came from the same or different populations. In this way the boundary lines between the unit areas may be accepted or rejected. The decision rests upon the probability of obtaining differences as great as those observed in the field from the same areal population, for even if the areal populations did not differ, samples drawn from each still would differ. How much must the samples differ for the population to differ? Objective procedures for this sort of decision making have been worked out by mathematical statisticians for statistical tests at various probabilities and degrees of freedom.

ORGANIZATION OF THE STUDY

The study in regional organization presented in this paper to illustrate the use of two statistical tests of significance for decision making may be stated as follows: In the organization of state groups West Virginia usually is placed with the Mid-Atlantic States (New York, New Jersey, Pennsylvania), or with the South Atlantic States (Maryland, Delaware, Virginia, North Carolina, South Carolina, Georgia, Florida). Why not place West Virginia with the East South Central States (Kentucky, Tennessee, Alabama, Mississippi)? The question can be answered statistically by comparing the distribution of the same trait among the three regions of state groups, or areal constructs, and West Virginia.

West Virginia should be placed in the state region with which its population of selected traits is homogeneous. The number of workers in 1950 in manufacturing and in the primary industries (agriculture, forestry, mining, fishing) were compared. This trait was selected because the relative number of workers in the primary and secondary sectors of the economy is a fundamental characteristic of an area. It reflects the way in which regional resources are utilized and is in turn related to problems of economic growth and development.

The distribution of the labor force by industrial groups is viewed here as a trait which is distributed areally among West Virginia, the Mid-Atlantic, the South Atlantic, and the East South Central States. Which of these four areas have different labor force populations; which are homogeneous? Table 1 presents these data. Mere inspection of the data can not answer these questions. Even if the counts were weighted to make them proportional to the total numbers reported, the investigator would have to make a decision as to the acceptable minimum differences in each area. Further, he would have to face the question of whether these reported differences may have been due to chance, i.e., differences which would be found among samples from homogeneous populations.

TABLE 1.—NUMBER OF WORKERS IN PRIMARY INDUSTRIES AND MANUFACTURING IN SELECTED STATES AND STATE REGIONS[1]

State and state regions	Number of workers, 1950, in— Primary industries	manufacturing
New Jersey	55,652	739,273
New York	184,556	1,772,707
Pennsylvania	356,515	1,395,166
Mid-Atlantic States	596,723	3,907,146
Delaware	11,517	41,075
Maryland	61,884	222,968
Virginia	203,063	235,501
North Carolina	366,515	408,904
South Carolina	199,314	210,748
Georgia	282,093	288,150
Florida	139,159	108,287
South Atlantic States	1,263,545	1,515,633
Alabama	280,960	225,126
Mississippi	307,901	90,336
Kentucky	315,578	150,774
Tennessee	263,016	239,446
East South Central States	1,167,455	705,682
West Virginia	196,101	118,525

[1] "Occupations and Industries in the Mid-Atlantic States," "Occupations and Industries in the South Atlantic States," and "Occupations and Industries in the South Central States," respectively, *Veterans Administration Pamphlets* 7–7.2, 7–7.5, and 7–7.6 (Washington, D.C., 1955).

It should be pointed out that complete counts were taken in the state unit areas and that the differences reported here are population differences. Thus any difference among the populations may be regarded as a real difference and not a sample difference. It is a moot question among statisticians whether whole finite populations of states or counties can be regarded as samples and submitted to significance tests. Much statistical research is conducted by sociologists on data of this type by assuming that the finite populations are samples from a larger universe. This hypothetical larger universe may be viewed as the infinite number of areal constructs with which it is possible to divide the space of the United States.[3]

While this objection may be raised against the validity of the present study, insofar as it employs significance tests, it is easy to visualize a situation in which it would not hold true. If an investigator were to regionalize an area according to some continuously distributed landscape feature and then collect data from each unit which amounted to a small percent (usually less than 20) of the finite population, objections to the testing of hypotheses of areal relationships for significance of difference on these grounds would be met. In any event, the testing operations are the same for each case.

Another objection often raised in statistical analyses of areal attributes has to do with the size and shape of the unit areas which may introduce errors. Randomness, or equality of opportunity of selection as a sample of each member of the population, is influenced by the varying sizes of the unit areas. Several methods of handling this problem have been suggested. These involve the use of an equal area sampling grid, area segments, areal weighting, average distance concept, or ignoring the problem.[4] Size of the sampling unit is of importance in correlation analysis because it is a technique for studying the relationship between quantitative characteristics distributed among units or unit areas of the same importance. In this study the relationship investigated is between non-quantitative (the regional frame) and quantitative characteristics (the data collected from the areas). Inequality of size is one attribute which results from the nature of the boundary lines and is reflected in the aggregate areal distribution of related landscape features. It is part of the regional matrix of features in which the investigator is interested. The peculiarities of regions are reflected not so much by the absolute values of particular features found in them, as they are by the density ratios of various traits.

The assumption of a hypothetical larger universe also permits the investigator to employ the entire finite population as a sample. Each sample thus has an equal opportunity of being selected because all are selected. While the area sampling problem may be avoided in regional analysis if data are collcted from all areal units, we raise the new problems mentioned above. One solution might be to use the computational operations of statistical tests as a measure of variation and difference without reference to probability statements.

The set of boundary lines used to create the areal frame by which the labor force traits have been sorted is political. Are such lines and the resultant areal patterns related to the traits studied? The geographic hypothesis is that they are. We wish to know more precisely which lines do separate homogeneous labor force populations, and which do not.

The political boundary lines of the states of the United States reflect the economic characteristics of the states. Why this is so is very complex in detail, but to a considerable extent variations in the economic history and the resource bases of different parts of the United States are associated with state boundaries. By grouping states together to form state regions the differences are accentuated. It thus may be possible to utilize these lines to set up regions of nonhomogeneously distributed features.

There is much to be said in favor of at-

[3] For a fuller treatment of these ideas, see M. J. Hagood and D. O. Price, *Statistics for Sociologists* (rev. ed.; Henry Holt & Co., 1952), pp. 284–94.

[4] W. C. Krumbein, "Statistical Analyses of Facies Maps," *The Journal of Geology*, Vol. 63 (1955), p. 452; A. H. Robinson, "The Necessity of Weighting Values in Correlation Analysis of Areal Data," *Annals*, Association of American Geographers, Vol. 46 (1956), p. 233; E. E. Houseman and T. J. Reed, *Applications of Probability Area Sampling to Farm Surveys*, Agricultural Handbook No. 67, United States Department of Agriculture (Washington, D.C., 1954.); H. H. McCarty, J. C. Hook, and D. S. Knos, with the assistance of G. R. Davies, *The Measurement of Association in Industrial Geography*, Department of Geography, State University of Iowa (Iowa City, 1956).

tempting to use states to regionalize the United States. There is a great deal of raw data on the social and economic characteristics of states and smaller political units collected by the U.S. Bureau of the Census which are available for study. To collect these data according to different areal patterns would be a formidable task, even if the alternative system could be agreed upon. The problem of assembling these data in a more functional spatial frame has occupied the Census Bureau.[5] The method followed is essentially that of procedure "2," described above. The Bureau has employed as many as 61 socio-economic indexes in constructing its pattern of economic regions and subregions. The delimitation procedure relies on a comparison of the index values among unit areas and unit area groups. Units are assigned to groups from which they are least deviant. The approach presented in this paper is markedly different.

PROCEDURE

In order to ascertain whether state lines can be used to regionalize the economy of the United States, the data on labor force distribution were assembled by the state regions indicated above and treated according to two tests of statistical significance, chi square and variance analysis. These tests are appropriate tools for the study of regions because they are applicable to problems dealing with an association between quantitative and non-quantitative characteristics. The areal construct is non-quantitative and the data collected from each unit area are quantitative. Is there a significant association between the two?

While chi square is a less efficient tool than variance analysis for decision making in regional construction, it is the most widely used non-parametric technique enabling the investigator to compare distributions without specifying their form. It is included here for illustrative and comparative purposes. The decisions reached by the two operations should agree rationally if we are to have objective procedures. It should be noted also that the *F* test, used here with variance analysis, is based on the distribution of the statistic *F*, which is generated as a ratio of two chi squares. This enables the investigator to make decisions about two sample variances and to introduce the question of internal regional homogeneity into the analysis.

The operational procedure involved in the application of the tests to this kind of spatial problem are discussed below.

Chi Square

Chi square is a method for directly comparing two distributions, in this study areal distributions. The two areal distributions which are compared are: (1) the distribution of the labor force actually observed among the state regions and (2) a labor force distribution calculated so as to be homogeneously distributed among the same state regions. The calculated population assumes that the state boundaries have no sorting influence and is weighted according to the total size of the labor force in each state region. Table 2 presents these data on the observed and computed labor force frequency distributions for each of the paired unit areas of state regions to be tested.

How do these two spatial distributions differ from each other? If they are samples from the same population, the boundary lines do not sort out the labor force. The hypothesis of independence between the state regions and their labor forces is accepted. If they are different, then the boundaries do result in a nonhomogeneous distribution and yield areal or regional differences. The hypothesis of independence is rejected. But by how much must they differ for the investigator to conclude that they are samples drawn from different populations (for even if they belonged to the same population the samples would differ because of chance variations)? This depends on the number of independent observations and the level of probability. For each operation there is a critical chi square value; if this is exceeded the differences between the two distributions are too great to be attributed to chance, and are regarded as statistically significant. The areal construct of state regions yields a spatial arrangement of the labor force distributions which are associated with the boundary lines in other than a chance manner. The regional structure has been validated.

If the critical chi square value is not exceeded the differences between the two dis-

[5] D. J. Bogue, *State Economic Areas*, United States Bureau of the Census (Washington, D.C., 1951).

TABLE 2.—OBSERVED AND COMPUTED LABOR FORCE FREQUENCY DISTRIBUTIONS FOR WEST VIRGINIA AND STATE REGIONS

Paired unit areas tested	Number of workers, 1950, in—				
	Primary industries		Manufacturing		Total
	Observed	Computed	Observed	Computed	
West Virginia	196,101	51,768	118,525	262,858	314,626
Mid-Atlantic States[1]	596,723	741,056	3,907,146	3,762,812	4,503,869
Total	792,824		4,025,671		4,818,495
West Virginia	196,101	148,439	118,525	166,186	314,626
South Atlantic States[2]	1,263,545	1,311,206	1,515,633	1,467,971	2,779,178
Total	1,459,646		1,634,158		3,093,804
West Virginia	196,101	196,095	118,525	118,531	314,626
East South Central States[3]	1,167,455	1,167,461	705,682	705,676	1,873,137
Total	1,363,556		824,207		2,187,763

[1] New York, New Jersey, Pennsylvania.
[2] Delaware, Maryland, Virginia, North Carolina, South Carolina, Georgia, Florida.
[3] Alabama, Mississippi, Kentucky, Tennessee.

tributions are still actual differences; but they are differences which may be due to chance, or differences which might occur between two samples drawn from the same population. If this is so, the state lines may not have sorted the labor force into a nonhomogeneous areally distributed population. On this account the regional construct is rejected. The tested statistical hypothesis of no relationship, the null hypothesis, between the boundary lines and the areal dispersion of the data is accepted.

The results of the chi square analyses on the tested pairs of regions indicated in Table 2 show that West Virginia and the Mid-Atlantic States, and West Virginia and the South Atlantic States reveal significant differences in the distributions by industrial groups of their labor forces as evaluated by this method. However, the differences between West Virginia and the East South Central States are insignificant. Their labor force populations are homogeneous; the differences observed may have been due to chance. The boundary lines between the unit areas have no sorting influence and West Virginia can be combined with the East South Central States because their labor forces are homogeneous.

With West Virginia placed in the East South Central States the geographic hypothesis to be tested is that the state regions, Mid-Atlantic, South Atlantic, and East South Central (including West Virginia), have labor force distributions which differ significantly from each other. The raw data may be assembled from Table 1, and the observed and computed distributions are given in Table 3. The results of the chi square test indicate that there is a significant difference among the three state regions.

The operations described above do not exhaust the possible manipulations of the data. But they have been carried far enough to indicate the usefulness and limitations of chi square in regional analysis. It appears to be an effective tool for assigning unknown unit areas to regions with which they are homogeneous. But it may not be sufficiently discriminating to evaluate several alternative regional constructs, especially when high frequencies are involved which may bias the test toward yielding significant results. For example, other groupings of states than that of Table 3 also yield significant chi square values.

Chi square is an over-all test of association. It does not measure internal variation among the state unit areas of which the state regions are composed. This could be accomplished by testing each state against every other state in its state region, or by testing each state against all the state groups. To a degree, this has been done for West Virginia. But to do this for all the unit areas would require many tedious operations. There is available, however, a procedure for evaluating inter- and intra-regional variation in one operation.

TABLE 3.—OBSERVED AND COMPUTED LABOR FORCE FREQUENCY DISTRIBUTIONS AMONG STATE REGIONS

State region	Number of workers, 1950, in—				
	Primary industries		Manufacturing		
	Observed	Computed	Observed	Computed	Total
Mid-Atlantic[1]	596,723	1,533,098	3,907,146	2,970,771	4,503,869
South Atlantic[2]	1,263,545	946,021	1,515,633	1,833,157	2,779,178
East South Central[3]	1,363,556	744,705	824,207	1,443,058	2,187,763
Total	3,223,824		6,246,986		9,470,810

[1] New York, New Jersey, Pennsylvania.
[2] Virginia, North Carolina, South Carolina, Georgia, Florida, Maryland, Delaware.
[3] Alabama, Mississippi, Kentucky, Tennessee, West Virginia.

Variance Analysis

When regions are constructed from smaller areal units and these are compared to each other for differences in traits, there are two sources of variation. There is the variation among the units within a region, and the variations among the regions. In building regions of homogeneous units, internal differences should be at a minimum and external differences at a maximum. Analysis of variance separates the two sources of variation, for it is a measure of association based on a ratio of between-region to within-region variations. The within-region variation is the variation of the areal units around their respective regional mean; the between-region variation is the variation of the regional means around the mean of all the regions. The ratio of the sums of these squared deviations, divided by their respective degrees of freedom, is the statistic F.

If the two mean square variances are equal, the value of F is one. The more F departs from one, the smaller is the probability that such a difference would be obtained by chance if the data were drawn from samples that came from the same population. If this probability is sufficiently low it can be concluded that the samples from which the within-region and the between-region mean square variances were calculated came from different populations.

If critical values of F have been exceeded at given levels of probability and degrees of freedom, the observed differences are significant, i.e., not due to chance. The regional construct being investigated may be said to yield an areal distribution of the tested characteristics which results in a significant association with the boundaries. It thus is possible to test the distribution of characteristics among several areal constructs and to arrive at a decision as to which one gives the maximum inter-regional differentials.

In this study there are three areal constructs which result when West Virginia is placed with the Mid-Atlantic States, with the South Atlantic States, and with the East South Central States. Ratios between the number of workers in 1950 in manufacturing and in the primary industries were calculated for each of the states from the data of Table 1. West Virginia then was placed successively with each state region and F tests of significance by the analysis of variance, with one criterion of classification and unequal size classes, were run separately on the three regional constructs. The results are given in Table 4.

In test A, West Virginia was placed with the East South Central States. The test result may be read as follows: The probability of obtaining an F value of 15.78 based on 2 and 12 degrees of freedom is less than .001. In test B West Virginia was placed with the South Atlantic States. The test result may be read as follows: The probability of obtaining an F value of 15.35 based on 2 and 12 degrees of freedom is less than .001. In test C West Virginia was placed with the Mid-Atlantic States. The results may be read as the probability of obtaining an F value of 5.17 with 2 and 12 degrees of freedom is more than .01 but less than .05.

What do these tests mean and how may they be interpreted to help provide an objective basis for the organization of regions? For each test the F value is a ratio of the within-region mean square variance to the between-region mean square variance. The test is to

TABLE 4.—ANALYSIS OF VARIANCE IN RATIOS OF NUMBER OF WORKERS IN MANUFACTURING AND IN THE PRIMARY INDUSTRIES, BY STATE REGIONS, 1950

Test	State region	Source of variation	Sum of squares	Degrees of freedom	Mean square variance	Ratio of variances F
A	1. Mid-Atlantic (N.Y., N.J., Pa.)	Total	199.08	14		
	2. South Atlantic (Md., Del., Va., N.C., S.C., Ga., Fla.)	Between class	144.26	2	72.13	15.78
	3. East South Central (*W.Va.*, Ky., Tenn., (Ala., Miss.)	Within class	54.82	12	4.57	
	$P\,[F_{2,\,12} = 15.78] < .001$					
B	1. Mid-Atlantic (N.Y., N.J., Pa.)	Total	199.08	14		
	2. South Atlantic (*W.Va.*, Md., Del., Va., N.C., S.C., Ga., Fla.)	Between class	143.09	2	71.55	15.35
	3. East South Central (Ky., Tenn., Ala., Miss.)	Within class	55.99	12	4.66	
	$P\,[F_{2,\,12} = 15.35] < .001$					
C	1. Mid-Atlantic (*W.Va.*, N.Y., N.J., Pa.)	Total	199.08	14		
	2. South Atlantic (Md., Del., Va., N.C., S.C., Ga., Fla.)	Between class	92.18	2	46.09	5.17
	3. East South Central (Ky., Tenn., Ala., Miss.)	Within class	106.90	12	8.91	
	$.01 < P\,[F_{2,\,12} = 5.17] < .05$					

ascertain if these two estimates of variance could be expected from the same population. The actual tested hypothesis is that the two estimates of variance, in test A for example, 72.13 and 4.57, might have been made from the same population, considering the degrees of freedom involved in each estimate. If they did come from the same universe, the F value would be one because the estimates of variance would be the same and the within-region variation would equal the between-region variation. But some variation from one can be expected to occur even if they were drawn from the same population. For values of F less than the critical value, the tested hypothesis is accepted, and there is no significant difference between the two estimates of variance. There is just as much variation within the state regions as there is between them. The regional construct is invalid. When the critical values of F are exceeded the two estimates of variance differ significantly, and the tested hypothesis is rejected. The areal construct has resulted in a spatial distribution of the tested data which significantly differentiates the internal regional variation from the external regional variation. There is a significant difference between the variation within the state regions and the variation between the state regions.

In tests A and B the between-region estimates of variance are 15.78 and 15.35, respectively, times as great as the within-region estimates of variance. The probability of obtaining such wide F ratios if the estimates were made on the same population is less than .001. It is a likely conclusion that the estimates were made on different populations. In test C, however, the between-region estimate of variance is 5.17 times as great as the

within-region estimate of variance. The probability of obtaining such an F ratio if the estimates were made on the same population is less than .05 but more than .01. This is too great a chance to conclude that the estimates came from different populations. Therefore we conclude that the differences in labor force distribution among state regions which occur when West Virginia is placed with the Mid-Atlantic States may be as great as the labor force differences within the state regions.

There is still another important aspect to the use of variance analysis as a tool in region building. In the tests various numerical values of F were obtained. Two of the values were great enough to provide significant results at the .001 level. Does this mean that the two regional constructs are equally valid, or can we make a further evaluation between them? If the same number of degrees of freedom is involved in the tests, as was the case in this study, the F value can be used to measure the relative degree of association between the regional construct and the spatial distribution of the labor force. The F value of 15.78 obtained in test A when West Virginia is placed with the East South Central States shows a higher degree of association than the F value of 15.35 obtained in test B when West Virginia is placed with the South Atlantic States. It thus is possible to rate several alternative regional constructs based upon quantitative expression which reflects their degrees of association with specified traits.

DISCUSSION

Two statistical tests of significance have been employed to evaluate the relationship between regional constructs and the areal dispersion of selected traits within that regional frame. The tests were conducted on the same geographical hypothesis. How do the results compare? The tests are not comparable statistically for each is made on a different statistical hypothesis. But do our rational interpretations of the two tests on the same problem lead to the same geographical conclusion?

In chi square, significant differences in the labor force distribution were obtained at .001 level when West Virginia was tested with the Mid-Atlantic States, but no significant difference was found between the labor force distributions of West Virginia and the East South Central States. It was concluded that West Virginia should be combined with the East South Central States.

In analysis of variance it was found that the widest ratio of the between-region to the within-region variances was obtained when West Virginia was placed with the East South Central States. The high F value was the result of a regional construct which gave minimum internal differences and maximum external differences. Thus chi square and variance analysis support the conclusion that the labor force distribution of West Virginia resembles that of the East South Central States more closely than the Mid-Atlantic or Southeastern States.

The use of significance tests as a decision making tool must be interpreted conservatively. Tests do not prove the existence of regions, for by varying the conditions of the tests with respect to levels of probability, number of independent observations, and even the kind of test itself, other regional patterns may be constructed statistically. By selecting different data other regions may be recognized. The investigator does not know if he has achieved the best classification, or if his procedures are converging on an optimum classification.[6] But he can answer the more limited questions: To what region should this particular unit area be assigned? And, does this particular set of boundary lines spatially array the given traits into different homogeneous areas?

The construction of state economic regions and other kinds of regions by this method also detects and considers the dynamic aspects of regionalism. As different resources come into use and are reflected in the social and economic charactristics of areas, it may be necessary to realign states or other unit areas. These changes must be of a sufficient magnitude to alter the test results.

The methods discussed in this paper contribute to the development of a more rigorous analysis of the regional process. They may be applied to more precise studies of the association between the natural environment and the socio-economic response pattern observed on the face of the earth. The concept of regions furnishes the basis of a methodology for studying the dynamics of the spatial interac-

[6] These points were raised by Dr. W. L. Garrison, University of Washington, in a personal communication.

tion of uniquely distributed traits. For when boundary lines are drawn to create a system of areal units a situation has been established in which it is feasible to test the relation between the basis of the lines and any aspect of the landscape. It is possible that the continued erection and testing of working hypotheses may lead to the development of a verifiable geographic theory of regionalism with predictive qualities.

THE SPREAD OF MINING IN THE COAL BASIN OF UPPER SILESIA AND NORTHERN MORAVIA

NORMAN J. G. POUNDS
Indiana University

THE coal field of Upper Silesia and Northern Moravia[1] is one of the richest in Europe. Since its wealth was first discovered in the eighteenth century, it has been shared by two or more countries. The coal field, however, forms a single geological unit, and it is the purpose of this paper to trace the spread of mining over the coal basin of Upper Silesia and Northern Moravia as a whole; to indicate the variations in quality and total thickness of the coal in different parts of the basin, and to show what influence these had on the expansion of mining itself.

Physical factors have not been the only important influences on the course of mining. The coal basin has, throughout its modern history, been divided politically. The legal, commercial, and industrial policies of each of the powers which have shared the coal basin have influenced the course of its development in their respective sectors. The political division of the area has hitherto prevented a rational and coordinated exploitation of the basin as a whole. Now the greater part of the coal basin is under Polish rule, and a more carefully planned exploitation of its resources is being undertaken. This seems, therefore, to be an appropriate time to examine the growth and spread of coal-mining in this area, to assess its importance, and to indicate, in the light of known resources, the probable course of its future expansion.

GEOLOGICAL FACTORS

Most of the coal basin is today covered with Secondary and Tertiary deposits. These are generally shallow, and the technical problems of piercing them are not great. Along the southern margin of the basin, however, the coal series was deeply buried beneath *Flysch* beds, formed in the course of the building of the Carpathian Mountains. Few bores have yet reached the coal measures on this side, and the limits of the basin here are not clearly known. Elsewhere, however, the behavior of the coal beds in both the exploited and the nonexploited parts of the basin is known with a considerable degree of precision.

The basin is triangular in shape (Fig. 1), with a superficial area of about 5400 square kilometers[2] (about 2085 square miles). Uncertainty regarding the southern margin of the basin introduces discrepancies between various estimates of its area. The structure of the whole is relatively simple. The coal measures themselves are grouped, on the evidence of rock type and of contained fossils, into a lower and an upper series. These are known respectively, from their relative positions in the basin, as the Marginal Group (*Grupa brzeżna, Randgruppe*) and Basin Group (*Grupa łękowa, Muldengruppe*). They differ from one another on both palaeontological and petrological grounds, and the coal produced is distinctive both in quality and usefulness. The lowermost part of the upper group is distinguished from the rest as the Saddle Beds (*Warstwy siodłowe, Sattelflöze*).[3] It contains the coal most nearly approximating a true coking coal to be found in the northern part of the basin.

The simple structure of the coal basin is interrupted by a number of folds. An anticline trends from west to east across the basin, dividing it into the small northern, or Bytom (Beuthen) Basin and the larger, southern or Main Basin. This anticline, known as the Saddle (*Siodła, Sattel*) brings the relatively valuable Saddle Beds to the surface, where they can be mined with ease, despite the considerable faulting which accompanied the formation of the folds. As it is continued eastward, the line of this main anticline is turned somewhat toward the southeast, and its amplitude

[1] The term "Moravia" is strictly a misnomer and is used here for that part of the coal basin which lay within the former Hapsburg provinces of Moravia and Austrian Silesia. The larger part of the coal production was always from the latter.

[2] Stefan Czarnocki, *Polskie Zagłębie Węglowe* (Warsaw, 1935), pp. 2, 177.

[3] The chief study of the structure of the basin is that of S. Czarnocki, *op. cit.* An earlier but valuable work is that of R. Michael, "Die Geologie des oberschlesischen Steinkohlenbezirkes," *Abhandlungen der königlichen preussischen geologischen Landesanstalt*, neue Folge, LXXI (1913).

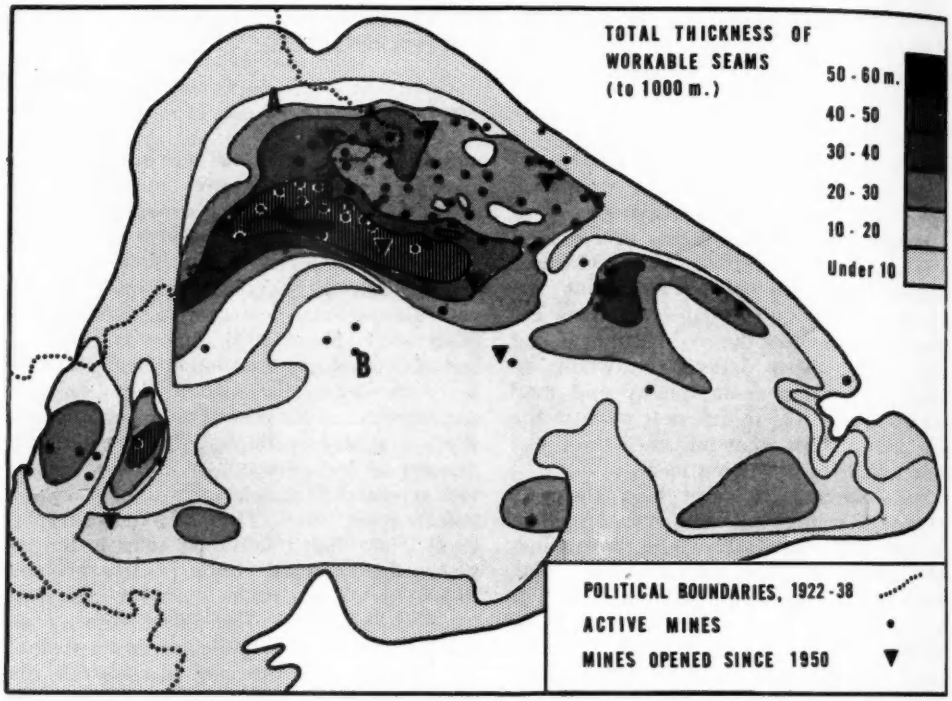


FIG. 1. Thickness of coal in the Upper Silesian–Moravian basin (after S. Czarnocki), and distribution of active mines.

becomes less. The Bytom Basin is characterized by small folds of similar direction, but the Main Basin is little disturbed, and the coal beds are in general level over the whole of its central part. Parallel with the western edge of the basin are north–south trending folds, constituting minor "saddles," and also a strongly developed line of faulting.

The quality and amount of coal varies not only vertically, in different parts of the coal series, but also horizontally. The lower or Marginal Group is relatively poor, and, it is estimated, contains little more than a tenth of the total reserves. The richest part of the whole coal series is the lower part of the upper, or Basin Group, where the Saddle Beds constitute a group of great value. Those parts of the coal basin where these beds come within the range of profitable mining have a quite exceptional importance.

Individual beds and groups vary greatly in thickness over the basin as a whole. As a general rule, seams are thickest to the north and west, but thin toward the south and east. This thinning of the seams is, however, accompanied by the attentuation of the intervening barren rocks. These may even disappear, so that seams merge with one another. In this way coal beds of exceptional thickness, such as a 10-metre seam, are formed in the northeast of the coal field.[4] The upper part of the coal series has been eroded from the Saddle, so that, though the Saddle Beds are thick, the total thickness of coal in this area is not great. The greatest thickness of workable coal lies to the south of the Saddle, where it amounts to as much as 60 metres of workable coal within 1000 metres of the surface. In the Bytom and Main basins, the total thickness of coal is considerably less than half of this. The only other part of the field where large combined thicknesses of coal are met with is to the west, where faulting has led to the preservation of a larger proportion of the total series. Figure 2 shows the variations in the combined

[4] Michael, *op. cit.*, p. 192, demonstrates the convergence of six seams in the Saddle Group to form the single Reden seam with 12.03 metres of coal.

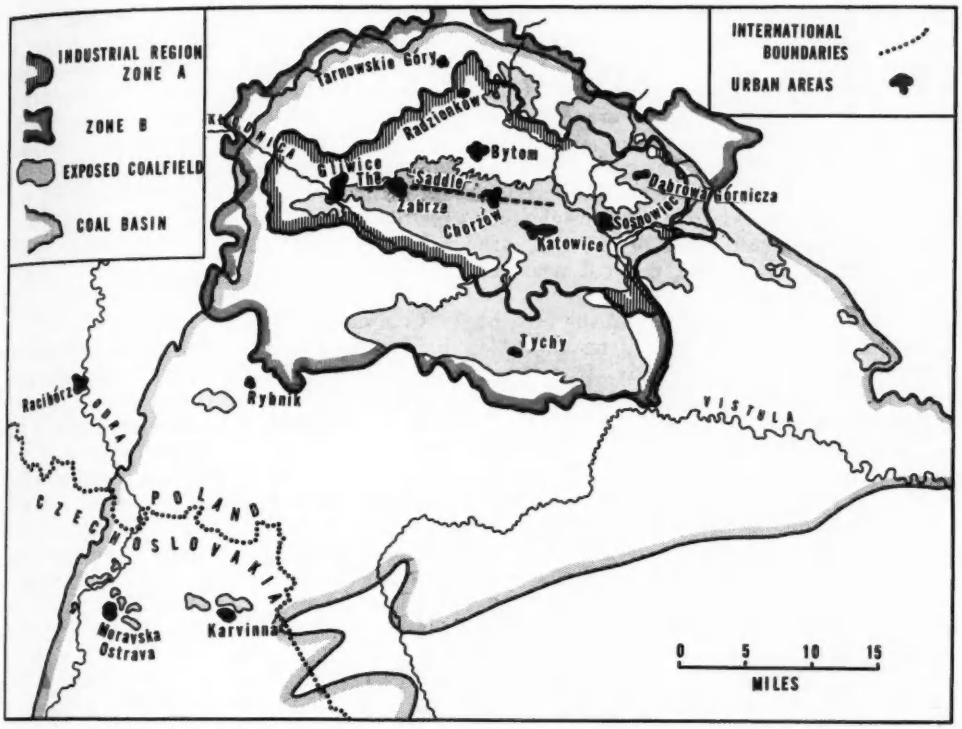


FIG. 2. The Upper Silesian Industrial Region, as defined by the Polish Government, 1953.

thickness of workable seams over the basin as a whole.

The quality of the coal varies over the basin, and appears to be related rather to the degree of folding to which it has been exposed than to the age and depth of the coal. All coals of the Upper Silesian–Moravian basin contain a relatively high proportion of volatile constituents. Coal with as little as 28 percent of gaseous content occurs only in the extreme southwest, within the present borders of Czechoslovakia. Elsewhere, the coal belongs to one of only two types: gas coal with 31 to 40 percent volatile matter, and flame coal with over 40 percent. This effectively removes it from the category of true coking coal.

The gaseous content of the coal increases to the east and south of the basin and is highest in the unfolded beds of the Main Basin. Coal with a relatively low volatile content is associated only with the anticlinal folding of the Saddle and with the more strongly folded of the beds of the west and southwest of the basin. True coking coals, similar in physical and chemical composition to the *Kokskohle* of the Ruhr, is thus absent from all areas except the strongly folded Ostrava area of Northern Moravia.[5] Elsewhere, coal with up to 36 percent gas is used for making metallurgical coke. This coke is, however, friable, and its use imposes irksome restrictions on the construction and operation of local blast-furnaces.[6] The improvement of the quality of metallurgical coke and the preparation of a firm coke from the weakly caking gas and gas–flame coals are urgent problems, and it is not improbable that some success will be achieved in the near future.

[5] On the relationship between folding and gas content, see S. Czarnocki, *op. cit.*, pp. 132–3.

[6] Franciszek Byrtus, "Baza węglowa Polski jako podstawa rozwoja przemysłu hutniczego," *Wiadomości Hutnicze*, X (1954), pp. 207–12. See comments on the quality of Polish coke in J. M. Montias, "The Polish Iron and Steel Industry," *The American Slavic and East European Review*, XVI (1957), pp. 311–12.

Most of the coal basin is hidden beneath later deposits. Triassic beds, predominantly limestone, have been eroded from all except the northern extremity of the field. Over much of the remainder of the area Tertiary and Quaternary deposits were laid down. Though most of the area was covered by the first advance of the ice sheet, this has left little trace, but diluvial deposits, originating from subsequent advances of the ice, are widely distributed. Indeed, the coal measures themselves outcrop over only about 300 square miles, some 15 percent of the area of the basin. Most significant by far of these outcrops is that of Upper Silesia. It stretches from Gliwice (*Gleiwitz*) eastward to Dąbrowa at the eastern margin of the basin, with an extension south into the Main Basin and north to the Bytom Basin. The significance of this outcrop lies in the fact that a large part of the Saddle is here exposed. Thus, the most valuable part of the coal-basin was exposed to prospectors from the earliest days of mining activity, and its development was encouraged by the growth of the Upper Silesian metallurgical industry.

Several small and relatively unimportant outcrops occur in the eastern extension of the coal field. In the southwest of the basin, partly within the present border of Moravia, are several small exposures of the coal field (Fig. 1). Though they cover only a very small area, the overlying beds which separate them from one another are thin and present no serious obstacle to mining.

Over most of the Main Basin the coal series is deeply buried beneath Secondary and Tertiary deposits. Shafts must here be sunk through very wet beds, and the profit to be obtained from the relatively thin seams of gas and flame coal has not been encouraging. In consequence, mining has not yet spread deeply into the Main Basin. The Bytom Basin, on the other hand, is smaller and less deep, and the quality of the coal higher. Mining spread over this part of the hidden field at a relatively early date.

POLITICAL FACTORS

Coal mining, if it was carried on at all, can have been of no more than slight, local importance before the second half of the eighteenth century. In 1741 the conquest of Upper Silesia by Frederick II gave a large part of the future coal field to Prussia. The rest was divided between the Austrian Hapsburgs and the Kingdom of Poland. The subsequent partitions of Poland divided the latter's share between Prussia and Austria. The pattern of political boundaries was again changed as a result of the creation by Napoleon of the short-lived Grand Duchy of Warsaw. With the absorption of the latter into the possessions of the Russian Tsar, the political division of the coal field (Fig. 6) which lasted for over a century, was created.

Although the mining of coal had begun before this threefold division of the coal field, the chief period of growth was during the nineteenth century, and it was then that the present distribution of mining was established. The course of this development was circumscribed by the pattern of political boundaries that had been imposed on the area. The rate of growth, technical equipment, and commer-

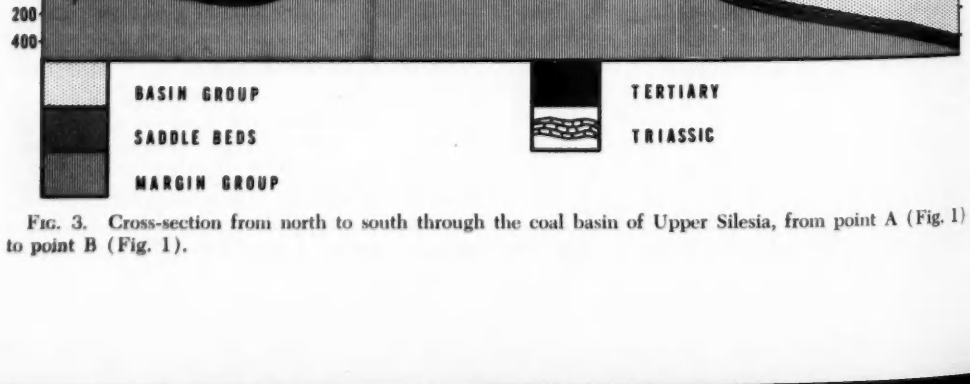


FIG. 3. Cross-section from north to south through the coal basin of Upper Silesia, from point A (Fig. 1) to point B (Fig. 1).

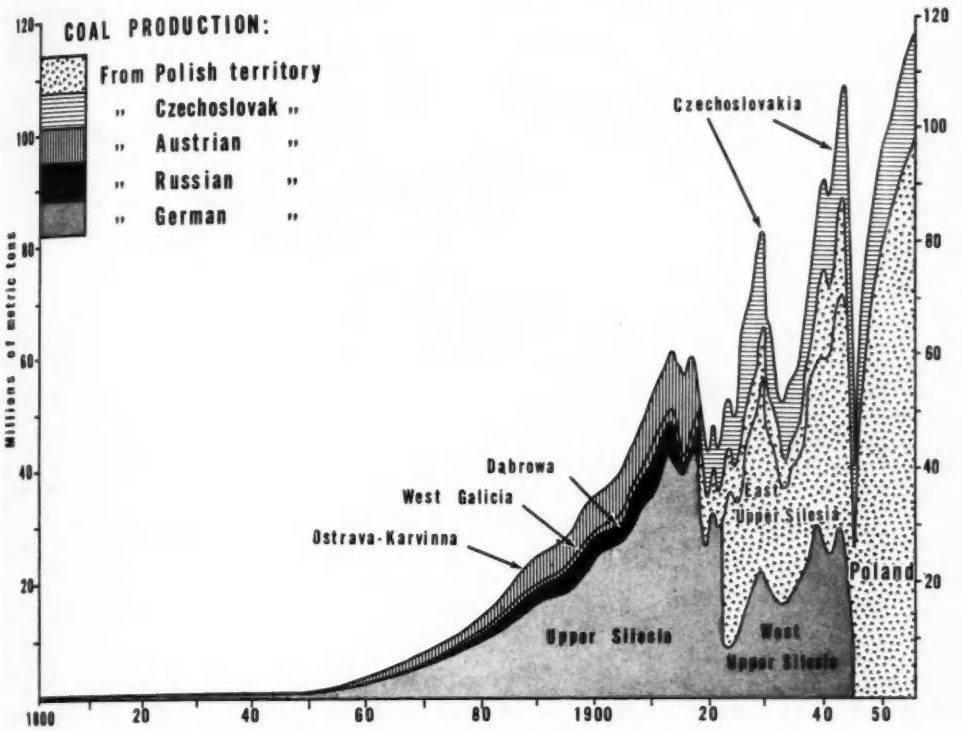


FIG. 4. Coal production from the Upper Silesian–Moravian basin, 1800–1955.

cial prospects differed sharply between the three political divisions of the area.

A consequence of Germany's defeat in 1918 and of the simultaneous creation of the Republics of Poland and Czechoslovakia was the establishment of a new political pattern. Moravia and, after prolonged debate, the southwestern part of Těšín (Cieszyn, Teschen) went to Czechoslovakia. Poland took the former Russian sector together with Austrian West Galicia. The German sector, after negotiation and armed rising had failed to settle the difficulty, was divided on the basis of the plebiscite, and the more easterly part was allocated to Poland.[7]

The confusion and disruption that resulted from this division are well known, though their ill consequences were, in fact, minimized by the international regime established for certain spheres of public administration by the League of Nations.[8] In anticipation of the defeat of Germany in the Second World War, the allied powers agreed at Teheran (1943), confirming their decision at Yalta (1945), to restore to Poland her long-lost western territories, as compensation for the loss of her eastern. As a result of these changes, the coal field was left divided unevenly between Poland and Czechoslovakia. The latter retained the sector which she had possessed before September, 1938; the rest became Polish.

GEOGRAPHICAL EXPANSION OF MINING

From its beginning in the eighteenth century, coal mining expanded both in the volume of coal produced and in the area mined. The graph in Figure 4, which shows the growth in output, demonstrates the relatively

[7] R. Hartshorne, "Geographic and Political Boundaries in Upper Silesia," *Annals*, Association of American Geographers, XXIII (1933), pp. 195–228; Gerhard Wende, "Die Auswirkungen der Grenzziehung auf die oberschlesische Montanindustrie," *Schriften des deutschen Ausland-Instituts Stuttgart*, Reihe E, Bd. 7 (1932), p. 3.

[8] George Kaeckenbeeck, *The International Experiment of Upper Silesia* (Oxford, 1942).

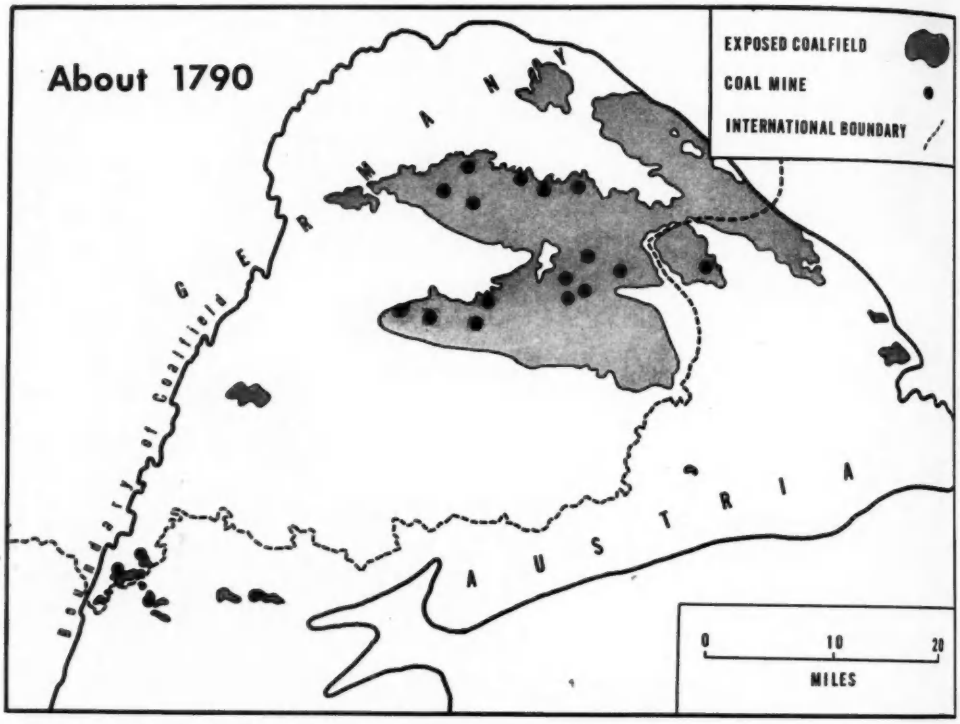


FIG. 5. Coal mines in the Upper Silesian–Moravian basin, about 1790. Boundaries shown as established at the Third Partition of Poland.

rapid growth of production in the German sector, due in part to the more developed German economy, in part to the use of water transport on the Odra (Oder) River. It shows a more modest growth in Moravia and a condition of near stagnation which characterized Russian-held Dąbrowa and Austrian-held West Galicia during much of the nineteenth century. This contrast is not explicable wholly in geographical terms; it results in large measure from political, social, and economic conditions within each of the three states which together divided and ruled the area of the coal field.

Inevitably coal-mining began on the exposed areas of the coal field. The earliest workings were in use by the middle of the eighteenth century in the territory of Poland that was to pass into Russian hands. Under the influence of the Prussian government, mining was then developed in German Upper Silesia, along the exposed Saddle. The exploitation of nearby deposits of iron, lead, and zinc ores created a demand for fuel, and coal-mining and metallurgy grew in close dependence upon one another.

In the German sector of the coal field there were by 1790 about fifteen mines active in the northern part of the coal basin, and at least two in the southwest (Fig. 5). All had been developed on the exposed area of the coal field, and there is no reason to suppose that the exposures were then known to be continuous with one another beneath the cover of later deposits.[9] Coal was discovered in northern Moravia in 1767,[10] and the first mine was

[9] *Jahrbuch des schlesischen Vereins für Berg- und Hüttenwesen*, I (1859), pp. 371–3; see also Hermann Fechner, "Geschichte des schlesischen Berg- und Hüttenwesens in der Zeit Friedrich's des Grossen, Friedrich Wilhelm's II und Friedrich Wilhelm's III, 1741 bis 1806," which appeared in nine installments in *Zeitschrift für das Berg-, Hütten- und Salinenwesen im preussischen Staate*, XLVIII (1900), IL (1901), and L (1902).

[10] R. Drapala, "Wie die Steinkohle zu Schles.-Ostrau wirklich entdeckt wurde," *Montanistische Rundschau*, XIX (1927), pp. 235–8; 259–65.

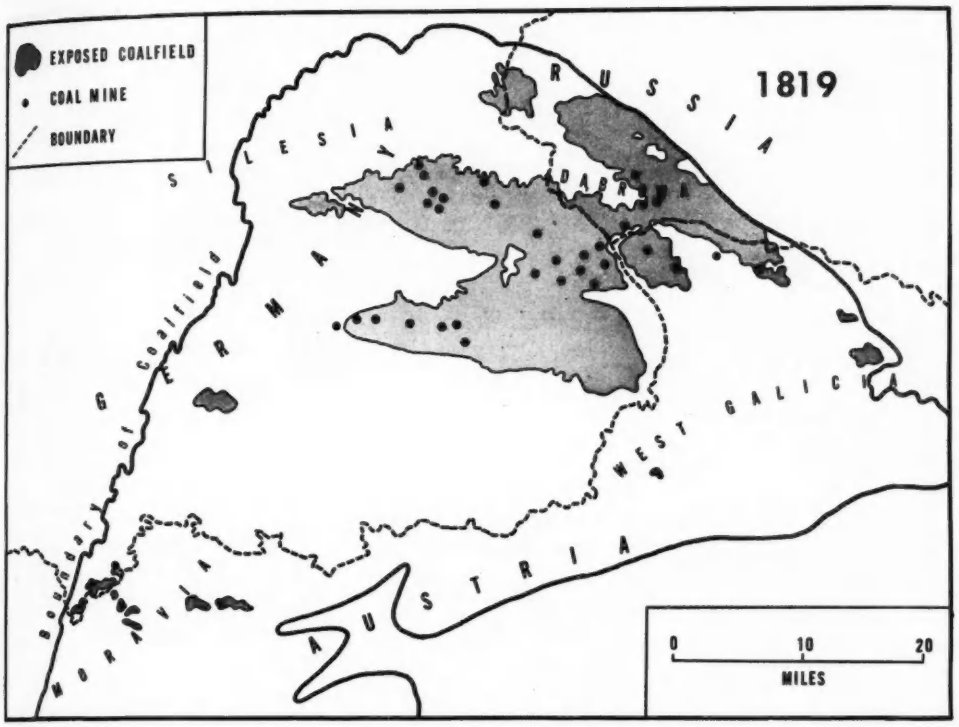


FIG. 6. Coal mines in the Upper Silesian–Moravian basin in 1819. Boundaries shown as established in 1815.

opened in 1790 at Karvinná.[11] Coal-mining had begun at least as early in the sector of the coal field that later came under Russian control, and in 1795 the earliest mine was sunk in Austrian-held Galicia.[12]

Von Oeynhausen's geological map of part of Upper Silesia, dated 1819, shows the active mines in the northern half of the basin,[13] in Germany, Russian-held Poland, and Austria (Fig. 6). They had doubled in number in the course of the previous thirty years, but still only two lay outside the area of the exposed coal field. Data are unsatisfactory for the mines in the southwestern part of the field at this date: we may assume that little progress had been made, but that mines active before 1800 were still in production.

The rate of economic growth in all sectors of the coal basin remained very low until after 1850. Then, with the coming of the railroads and the expansion of metallurgical industries, expansion was more rapid. Figure 7 shows the distribution of mining activity about 1860. The positions of shafts is taken from Carl Mauve's map,[14] but the representation of a mine here does not necessarily mean that it was in actual use. The totals, in metric tons, are for each of the Prussian mining districts (*Bergreviere*).[15] The output of individual mines is not known, and the positions of the ten mines in the Ratibor (Racibórz) district

[11] "Ueber die geognostischen Verhältnisse und den Bergbau des Orlau-Karviner Steinkohlenrevieres in Oesterreichisch-Schlesien," *Berg- und hüttenmännisches Jahrbuch*, XXI (1873), pp. 118–66.

[12] Franz Bartonec, "Die Steinkohlenablagerung West-galiziens und deren volkswirtschaftliche Bedeutung," *Oesterreichische Zeitschrift für Berg- und Hüttenwesen*, XLIX (1901), pp. 321–5; 336–40.

[13] Carl von Oeynhausen, "Geognostische Carte von Ober-Schlesien," published as folded sheets in von Oeynhausen, *Versuch einer geognostischen Beschreibung von Oberschlesien* (Essen, 1882). This work is the earliest which, to my knowledge, attempts to classify the types of coal found in this basin.

[14] Carl Mauve, *Flötz-Karte des Steinkohlen-Gebirges bei Beuthen, Gleiwitz, Myslowitz und Nikolai in Ober-Schlesien* (Breslau, 1806).

[15] *Jahrbuch des schlesischen Vereins für Berghau Hüttenindustrie*, I (1859), p. 107.

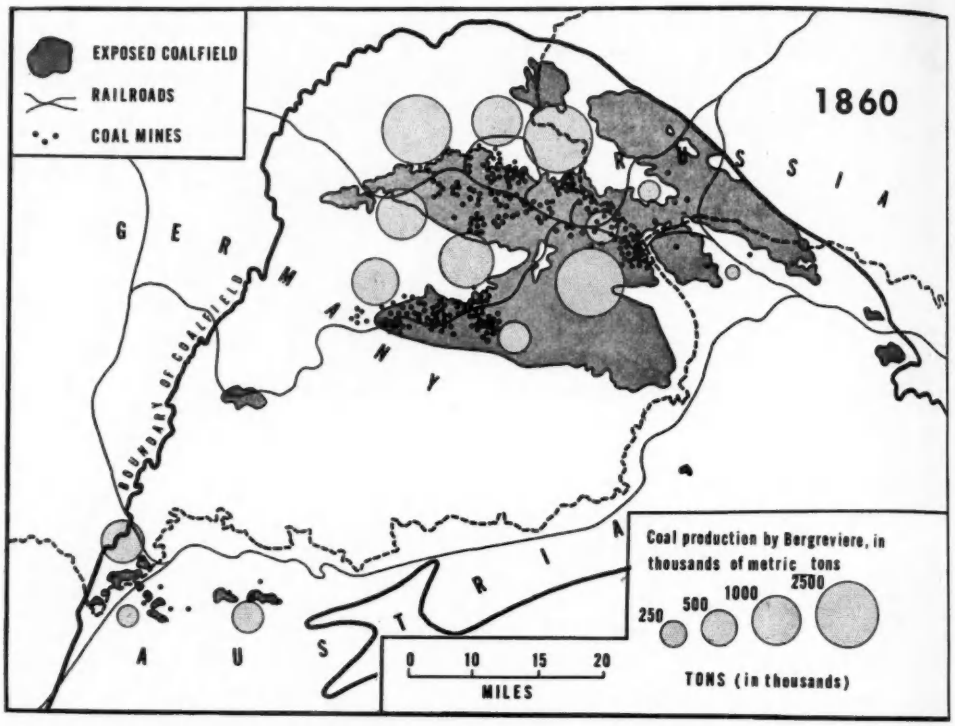


FIG. 7. Coal mines and coal production in the Upper Silesian–Moravian basin, 1860. Precise boundaries of the *Bergreviere* have not been determined. Boundaries shown as established in 1815.

are not given on Mauve's map. The positions of the Moravian mines are based on Balling's map of 1868.[16] The earlier pattern of mines in Dąbrowa and Galicia has been repeated on this map; though it had undoubtedly changed in detail, data are inadequate to show these changes with precision. A marked feature of the distribution of coal mining about 1868 is that it still keeps closely to the pattern established over half a century earlier; no shaft lies at a greater distance than about a mile beyond the margin of the exposed coal field. That this was in no way due to lack of geological knowledge is made clear by the detailed geological study of Roemer published only a few years later.[17] It is more likely that the geographical spread of mining was strongly influenced by the course of railroad building, and that the lack of a railroad was responsible for the absence of mines. On the map of coal-mining in 1860 and on maps for subsequent dates, the contemporary railroads are shown.[18]

The period from 1860 until about 1885 was one of steady growth, during which the output of the coal field as a whole increased almost threefold. This growth was accompanied by an actual reduction in the number of mines. Many mines and shafts were abandoned and production was concentrated at a smaller number of mines, each of very much greater size. Their distribution and output are shown in Figure 8.[19] An actual contraction

[16] Carl A. M. Balling, "Eisenindustrie-Karte von Mahren und Schlesien," *Berg- und hüttenmännisches Jahrbuch*, XVIII (1869). Mines are also listed in "Die frühen K. K. Kohlenwerke bei Mahrisch-Ostrau, *Oesterreichische Zeitschrift für Berg- und Hüttenwesen*, V (1857), pp. 27–9; 35–7.

[17] Ferdinand Roemer, *Geologie von Oberschlesien* (Breslau, 1870).

[18] Based on the admirable work of Marian Frank, "Przyczynek do badania wydobycia węgla w Polsce na tle rozwoju sieci kolejowej w Sląsko-dąbrowsko-krakowskim Zagłębiu węglowym do pierwszej wojny swiatowej," *Zeszyty Naukowe Wyższej Szkoły Ekonomicznej w Katowicach*, I (1957), pp. 1–36.

[19] *Zeitschrift des oberschlesischen berg- und hüttenmännischen Vereins*, XXVI (1887), p. 460.

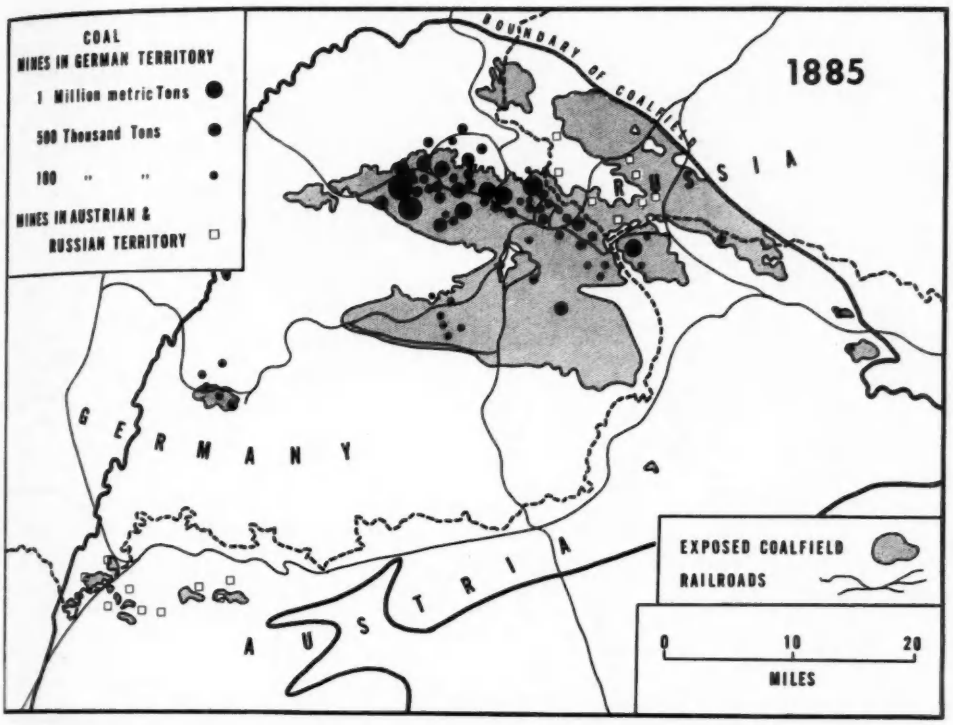


FIG. 8. Coal mines and coal production in the Upper Silesian–Moravian basin, 1885. Detailed statistics available only for the German and part of the Austrian sectors of the coal field. Boundaries shown as established in 1815.

can be seen in the area previously mined, accompanied by an expansion into the Bytom Basin and the opening up of mines in the Rybnik district to the west of the Main Basin, where the coal measures are not exposed at all. One may perhaps detect in this a search for coal of higher quality: the areas newly opened up produce some of the better qualities of coking-coal; those abandoned yielded only gas and flame coal.

In Dąbrowa there was also a sharp growth in coal production, which can unfortunately not be shown with the same precision as that of the German sector.[20] In West Galicia, the long-established mines at Jaworzno and Sierza expanded their output, and coal was obtained intermittently from a number of other mines.[21] There is no evidence for any significant change in the geographical pattern of mining in northern Moravia (see Figs. 8 and 9).

The map of coal-mining for about 1900 (Fig. 9) shows no marked change from that of fifteen years earlier.[22] Total production had almost doubled, but this increase was achieved by increased output from existing mines. It is not until the eve of the First World War that we find any marked change in the geographical pattern. Hitherto the number of mines had been contracting; now, in the map for 1912 (Fig. 10), we discern an increase. This is especially marked in the

[20] Kurt Flegel, *Die wirtschaftliche Bedeutung der Montanindustrie Russlands und Polens* (Osteuropa Institut, Breslau, 1920), p. 14.

[21] The Austrian statistics are quite unsatisfactory; see Hermann, "Ueber den Bergbau im Kreise Chrzanow in Galizien," *Zeitschrift des oberschlesischen Berg- und Hüttenmännischen Vereins*, XXXVI (1897), pp. 16–23.

[22] *Zeitschrift für das berg-, hütten- und salinenwesen im preussischen Staate*, IL (1901), Statistischer Teil.

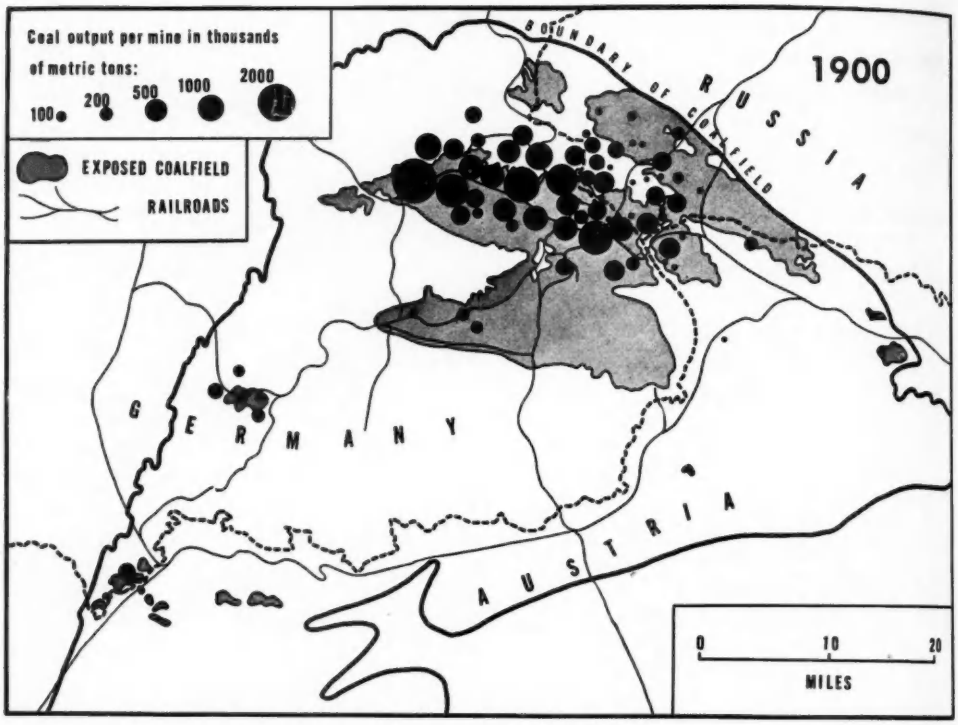


FIG. 9. Coal mines and coal production in the Upper Silesian–Moravian basin, 1900. Statistics not available for Moravia. Boundaries shown as established in 1815.

north, where an expansion is shown into the hidden coal field of the Bytom Basin, and near the western margin of the coal basin, where coking coal could be obtained.[23] This map also shows in greater detail than earlier maps the output of mines in Dąbrowa,[24] West Galicia,[25] and northern Moravia.[26] It is now possible to observe the spread of coal-mining on a large scale out over the hidden coal field, both north of the main outcrop and along the western margin of the basin.

After the period of uncertainty which followed the First World War, a new pattern of political boundaries was established. The coal-basin, of which the greater part had previously been held by Germany and Austria, went largely to Poland (see Table 1). The

TABLE 1.—CHANGE IN POLITICAL DISTRIBUTION OF THE COAL FIELD, 1914–1945[1]

Country	Percent of Coal Field Held in—		
	1914	1923	1945
Germany	48	10	—
Russia	8	—	—
Austria	44	—	—
Poland	—	72	82
Czechoslovakia	—	18	18

[1] Based on R. Michael, *Abhandlungen des königlichen preussischen geologischen Landesanstalt*, neue Folge, LXXI (1913), and S. Czarnocki, *Polskie Zagłębie Węglowe* (Warsaw, 1935).

[23] Gerke," "Ueber den Plesser Bergbau und die Boerschachte," *Zeitschrift des oberschlesischen berg- und hüttenmännischen Vereins*, LXVIII (1929), pp. 124–9.

[24] *Handbuch von Polen*, ed. E. Wunderlich (Berlin, 1918), pp. 419–40; K. Flegel, *op. cit.*; R. Michael, "Uebersichtskarte der Besitzverhältnisse im oberschlesischen Steinkohlenrevier und in den Nachbezirken," *Zeitschrift des oberschlesischen berg- und hüttenmännischen Vereins*, LII (1913), pp. 230–7.

[25] "Das galizische Kohlenrevier," *Montanistische Rundschau*, X (1918), p. 639.

[26] R. Michael, "Geologie des Oberschlesischen Steinkohlenbezirkes," *Abhandlungen der Königlichen Preussischen Geologischen Landesanstalt*, neue Folge, LXXI (1913); "Karte des Ostrau–Karwin–Krakauer Steinkohlenrevieres," *The Coal Resources of the World* (International Geological Congress Toronto, 1913), Map 44.

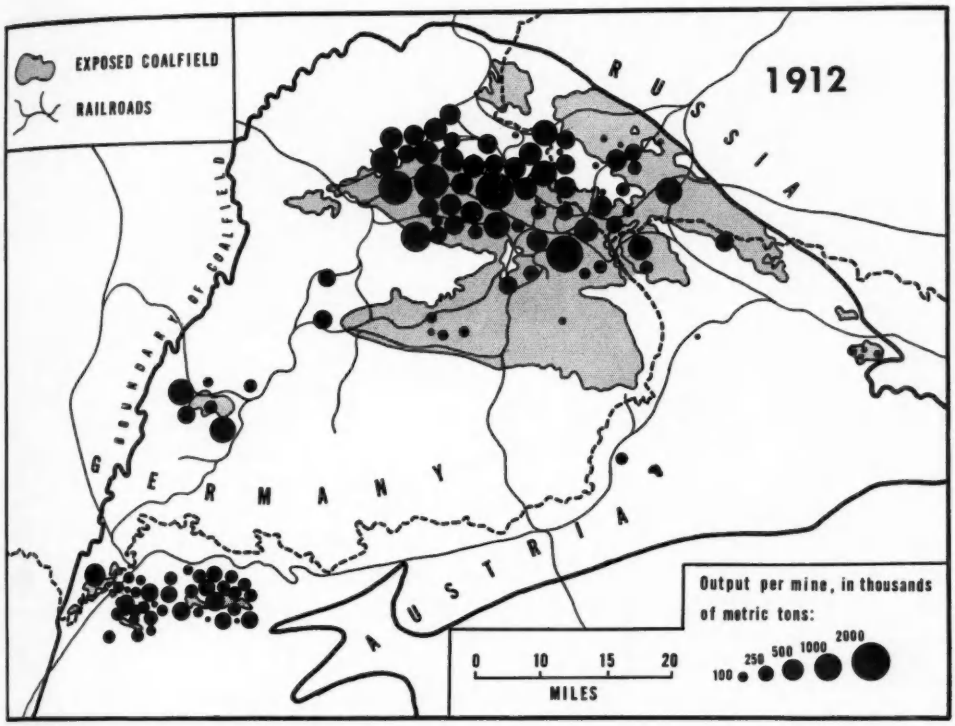


FIG. 10. Coal mines and coal production in the Upper Silesian–Moravian basin in 1912. Boundaries shown as established in 1815.

decade from the end of the First World War to the beginning of the Depression was characterized by a sharp increase in production from the small area of the coal field that remained in German hands and by a rationalization process in the large sector that had passed to Poland. Small mines in the former Russian and Galician sectors, which had hitherto enjoyed a measure of tariff protection, now closed or merged with their neighbors. At the same time, a German desire to dispense with imports of Polish coal brought about a considerable and perhaps unreasonable expansion in the German sector.

The new pattern is shown in Figure 11. It is marked by the disappearance of many of the small mines which had formerly existed in the Dąbrowa and West Galician sectors[27] and by the continued growth in the importance of mines in the west of the basin, especially near Rybnik.[28] In all sectors of the coal field, mechanization of the mining processes made considerable progress, and the expanded output of these years was in general achieved by a higher productivity of labor.

The pattern established by 1929 changed little during the next twenty years. During the period of German occupation of Poland, the mines were recklessly exploited, but new mines were not opened. Immense damage was done both to the mine workings and to the ground surface by the failure of the Germans to take proper steps to secure the roof and to prevent subsidence.[29] Nevertheless, an output of 91,930,000 metric tons was achieved in 1943 from the Polish and German sectors alone.

[27] *Zeitschrift des oberschlesischen berg- und hüttenmännischen Vereins*, LXVII (1928), pp. 268–9.

[28] Rudolf Wachsmann, "Der Steinkohlenbergbau im Rybniker Revier," *Zeitschrift des oberschlesischen berg- und hüttenmännischen Vereins*, LXIII (1924), pp. 233–8.

[29] Alfred Hornig, "Przemysł węglowy w Polsce podczas wojny 1939–1945," Dissertation in Typescript (Chorzów, 1948).

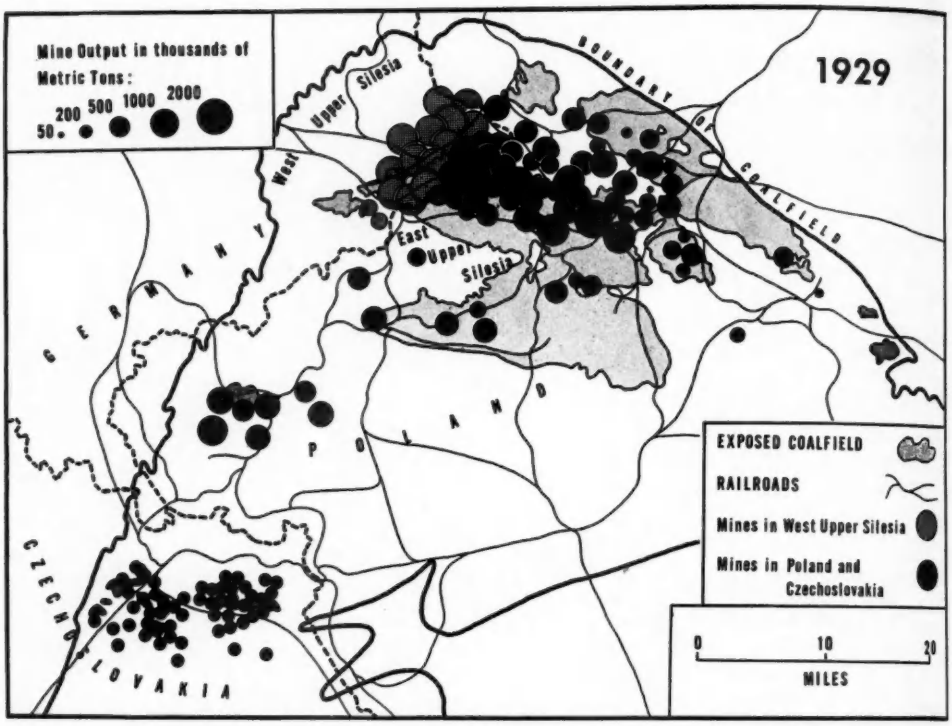


FIG. 11. Coal mines and coal production in the Upper Silesian–Moravian basin, 1929. The total output from the Moravian sector has been divided between the mines known to have been operating at this date. Boundaries shown as established in 1919–22.

The postwar period of planned development has brought about certain changes in the geographical pattern of coal-mining. The industrial expansion of Poland, as well as her overseas trade, is dependent upon an expanded coal production. The unified political control of most of the basin has indeed allowed a more rational development of the coal resources, and new mines and their ancillary services can now be located without reference to political boundaries. This is a great advantage and may be expected to increase the efficiency and productivity of the mines.

So far as can be ascertained from field work and such published sources as are available, Figure 12 represents the present distribution of mining activities in the coal basin as a whole. Some strip mines were opened in the Dąbrowa sector, but the dip of the beds makes any extensive use of strip mining impracticable. Their yield was small and they were regarded as only temporary, designed to maintain the level of production until new deep mines could come into production. It was anticipated that new mines would by 1955 contribute 9 percent of the total output, while new levels cut in existing mines would add a further 14.5 percent.[30] This hope was not realized. Work on new mines proceeded more slowly than had been expected, and the output of new mines was less than half of what had been hoped for.[31]

A Six-Year Plan provided for eleven new mines on virgin sites. Most of these appear to be in production, and their output is likely to increase gradually to the planned totals.[32]

[30] J. Kokot, "Przebudowa gospodarcza Śląska," in *Górny Śląsk: prace i materiały geograficzne* (Kraków, 1956), pp. 481–4.

[31] J. Dzierżynski, "Zadania przemysłu węglowego w pierwszym roku planu pięcioletniego," *Gospodarka Planowa*, XI (1955), No. 2, pp. 1–5.

[32] T. Muszkiet, "Główne problemy rozwoju górnictwa węglowego w planie 5-letnim," *Gospodarka Planowa*, X (1955), No. 4, pp. 1–8.

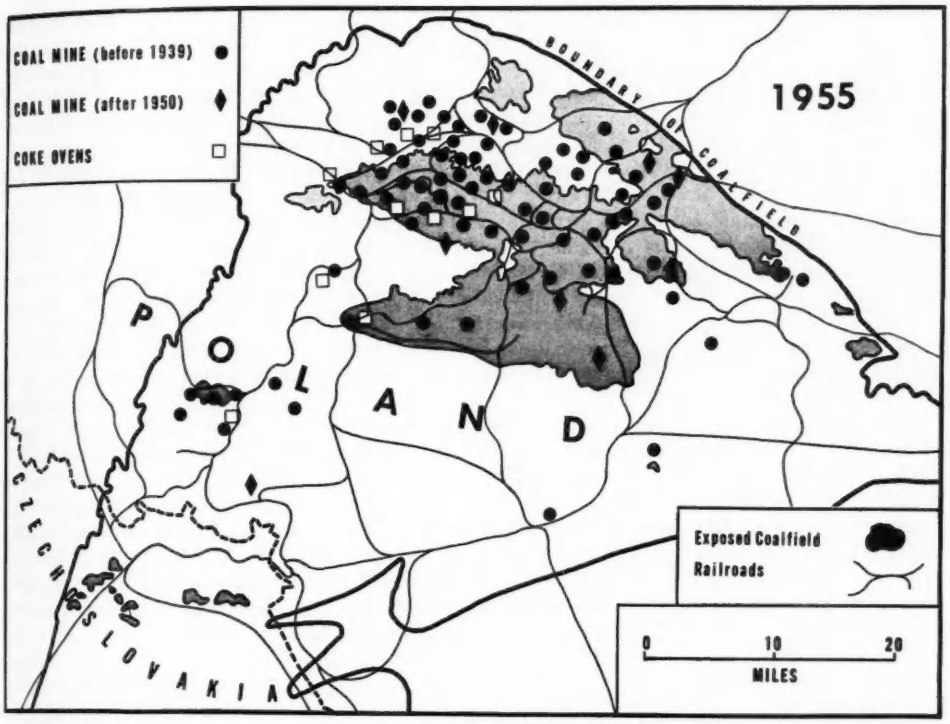


FIG. 12. Distribution of coal mines and coke ovens in the Upper Silesian–Moravian basin. Details of the active mines in the Czechoslovak sector are not available, but their geographical pattern is likely to be similar to that for 1929. Boundaries shown as established in 1945.

The distribution of these new mines is in keeping with trends already described. Two are in the hidden area of the Bytom Basin, and one is south of the Saddle. A new mine has been opened in the Rybnik area, and the rest lie in hitherto neglected areas toward the east of the basin. Two factors appear to have determined the location of new mines: (1) the availability of coking coal, whose distribution is very roughly shown by the occurrence of coke ovens (Fig. 12), and (2) the greatest thickness of workable coal. This latter correlation is demonstrated by a comparison of Figures 2 and 12. The gas and flame coal from some of the more easterly mines is destined for industry, power plants, and export. The coking coal is used in the blast-furnaces of Upper Silesia, Częstochowa, and Nowa Huta, but has to be supplemented by fuel from Lower Silesia and Moravia.

No comparable expansion has taken place in Moravia. The output of coal has increased in recent years, but this appears to have been due primarily to increased mechanization of the mines.[33] It is planned to open six new mines in the Ostrava–Karvinná area, and work has already been begun on some. These mines are not shown on Figure 12.[34] In view of the acute shortage of coking coal within the East European countries it is somewhat surprising that a more vigorous development of mining in northern Moravia has not taken place earlier.

TECHNICAL CONSIDERATIONS

It has been noted earlier that coal seams in the Upper Silesian–Moravian basin are in general relatively thick and that they are folded and faulted to only a slight degree. Exceptions to this occur in the Saddle and in north-

[33] *Prague News Letter*, October 29, 1955.
[34] *Ibid.*, March 2, 1957.

ern Moravia, but over the basin as a whole mining is technically relatively easy. At the same time, mines are relatively free of gas, and only in the western and southwestern parts of the basin are normal precautions taken against gas explosion.

Such easy technical conditions favor mining, and the Polish authorities are today taking advantage of them to expand production rapidly. Explosives are used in some of the thick seams to break down the coal-face, and in such workings the miner has only to load the coal on to the conveyor-belt. Where such rapid mining methods are not practical, coal-cutting machinery is slowly replacing the older methods of working.

These conditions which favor mining also set limitations to it. It is difficult, if not impossible, to support the roof in the workings by traditional timbering methods, and the mining of thick seams yields very little waste material for filling mined-out areas. The consequence has been excessive subsidence in rural areas and the cessation of mining at the margins of cities and of other settled areas. The *podsadzka* method of filling was introduced to meet this problem a half century ago, but only since Upper Silesia passed exclusively into Polish possession have the conditions existed for its use in most mines. Briefly, the worked-out space is enclosed with wooden boards and filled with sand and gravel, washed down from the surface by pipe.[35] As the filling dries it compacts somewhat, but it nevertheless in large measure cushions the roof collapse. The widespread adoption of *podsadzka*, along with the mechanization of coal-cutting and transport, is the chief means of increasing the output of older mines. It allows a larger proportion of the thick seams to be extracted, and it permits the coal to be taken out from beneath built-up areas.[36]

Use of the process, however, necessitates a large capital investment, not only in the concrete coffers, in which sand is stored and from which it is hosed down into the mines, but also in the railroad system which distributes the sand from the alluvial deposits along the valleys. The chief source is the Pustynia Błędowska, a large area of sand and gravel in the valley of the Biała Przemsza.[37] Other deposits are used in the Kłodnica and Czarna Przemsza valleys. Already, however, very large areas have been made agriculturally unproductive as a result of subsidence.[38] The use of sand is a palliative; it will reduce subsidence to small proportions, though at the expense of going ever farther afield for deposits of sand. Nor does it obviate the need for timbering. Indeed, the supply of mine timber suitable for the thick seams has posed a problem.[39]

The most serious deficiency in the coal basin is the scarcity of coking-coal. Northern Moravia produced coal with a volatile content of less than 30 percent, and coke made from it is used in Czechoslovakia as well as in Poland and the Danubian countries. Certain of the seams in the Rybnik area and in the western part of the Saddle yield a coal that can be used for coking, but the metallurgical industry suffers because of deficiencies in the fuel. The rest is gas and flame coal. It is used locally in the zinc-furnaces, still predominantly of the old, extravagant muffle pattern; in the thermal generating stations; and to an increasing degree in the chemical works. But there is a large export of coal from all sectors of the coal field, not only to other parts of Poland and Czechoslovakia, but also abroad.

The Czechoslovak export from northern Moravia appears to consist mainly of coking coal and to be more than offset by an import of coal of other types from Poland. The Polish export is on a larger scale and is an essential feature of the national economy. Since the re-establishment of the Polish state in 1918, coal has been a leading export. Before the war, the Scandinavian and Balkan countries were the chief consumers. Now about a quarter of the coal produced (24,146,000 metric tons in 1955) is exported, a third of it to the Soviet

---

[35] *Wiadomości Górnicze*, IV (1953), No. 10 was devoted to articles on the *podsadzka* process.

[36] *Życie Gospodarcze*, December 8, 1957, notes that the Katowice mine has been given a new lease on life as it can now extract coal from beneath the Huta Ferrum steel-works.

[37] Tadeusz Paszta, "Centralne zaopatrzenie w piasek podsadzkowy kopalń Śląskiego Zagłębia Węglowego," *Wiadomości Górnicze*, V (1954), pp. 207–10.

[38] Tadeusz Mrzygłod, "Potrzebna jest śmiała koncepcja deglomeracji GOP," *Życie Gospodarcze*, November 24, 1957; see also Alfred Hornig, "Formy powierzchni ziemi stworzone przez człowieka na obszarze wyżyny śląskiej," in *Górny Śląsk: prace i materiały geograficzne* (Kraków, 1955), pp. 123–49.

[39] "Jak zaoszczedzić drewno kopalniana?" *Wiadomości Górnicze*, IV (1953), pp. 139–42.

Union and most of the remainder to Czechoslovakia, Finland, Austria, and Sweden. In addition, coke is exported (2,240,000 metric tons in 1955) to East Germany, the Soviet Union, Hungary, and Romania.[40]

Coal is one of the few commodities with which Poland can pay for her imports, and the extension of her markets in the free world is a necessity if she is to develop her industries and improve her standard of life. Fears have been expressed[41] that American shipments to Europe may restrict Polish sales in the West European countries. To meet this threat, Polish coal producers were urged to give more attention to the quality and careful sorting of the coal exported. This, in turn, necessitates investment in coal-washing plants which have hitherto not been particularly prominent in Upper Silesia.

This paper is not concerned with the problems in social geography of the coal field. It may suffice to say that in the long-established mining areas there are extreme problems of congestion, made worse by the current expansion of the mining industry. At the new mines there is a housing shortage which makes the recruitment of labor difficult. The Upper Silesian industrial region (Górno-Śląski Okręg Przemysłowy) was constituted in 1953 as a planning area. It was divided into Zone A (Fig. 1), which included the heavily industrialized area and also the older mining centers, and Zone B, which is more lightly industrialized.[42] Zone A contains the thickest and most valuable coal deposits, and there the extraction of coal has the highest priority. Only those industries heavily dependent on coal may now be located there, and every effort is being made to move population out into the surrounding Zone B. Coal-mining is also important over almost the whole of the latter. New industrial plants are being located mainly in Zone B, and several dormitory towns have been established to house the workers not only of the new mines and factories, but also of the older industrial core of Zone A.

CONCLUSION

The Upper Silesian–Moravian coal field as a whole is probably the largest in Europe in both area and total reserves. It challenges comparison with the Ruhr but differs from it in (1) the absence of anthracite and the small proportion of coking-coal, (2) the relative lack of waterborne transport, and (3) the smaller scale of the industrial development that has taken place on it. To call it the "Eastern Ruhr"[43] is to exaggerate its industrial (as distinct from mining) role. It is, however, a source of energy for Poland, Czechoslovakia, and the Danubian lands, and if the system of canals currently planned between the Odra, Vistula and Danube is ever completed, the coal-basin may become a source of industrial power for a large area of East-Central Europe.

[40] *Rocznik Statystyczny 1956*, Głowny Urząd Statystyczny (Warsaw, 1956).

[41] *Życie Gospodarcze*, December 8, 1957.

[42] Janusz Ziołkowski, "Z problematyki przestrzennej i gospodarczej Górnośląskiego okręgu przemysłowego," *Przegląd Zachodni*, XIII (1957), pp. 257–79.

[43] P. H. Seraphim, *Industrie-Kombinat Oberschlesien: das Ruhrgebiet des Ostens* (Köln–Braunsfeld, 1953).

LETTERS AND COMMENTS ON *ANNALS* TOPICS

Letters and comments on *Annals* topics are welcome. Statements should be concise, restrained, helpful, and deal with *Annals* subject matter. The Editor reserves the right to accept or reject material out of hand and to make such changes as seem necessary or becoming. This feature will be run as often as available material warrants.

THE EDITOR

CHI SQUARE AS A TOOL FOR REGIONAL STUDIES

J. ROSS MACKAY
The University of British Columbia

In a recent interesting article by L. Zobler in the *Annals*, chi square was used for testing regional boundaries.[1] Although the use of chi square appears to be excellent for certain purposes of regional analysis, as Zobler has ably demonstrated, the validity of the chi square test as employed in part of the article is open to question on theoretical grounds. The test is based upon the derivation of chi square as follows:

$$\chi^2 = \Sigma \left\{ \frac{(f_o - f)^2}{f} \right\}$$

where f_o is the observed frequency and f the theoretical or expected frequency.

In the application of the chi square test, the frequencies used must be absolute, not relative. This places a limiting condition upon the use of the chi square tool. "The reason for this condition is obvious: The significance of a given divergence of f_o from f depends on the absolute magnitude of f. The divergence of 4 from 3 may be negligible, the divergence of 400 from 300, which is the same in relative terms, may be highly significant."[2] What, however, is the absolute frequency of a geographic area? Should it be measured in acres, hectares, square kilometers, or square miles? Should the absolute frequency of a geographic area vary with the unit of mensuration? For instance, if the raw data for a study were collected upon the basis of acres, there would be one set of frequencies; if upon the basis of hectares, there would be another; and so forth. Of course the relative sizes (frequencies) would be the same irrespective of the units of mensuration employed, but the theory of chi square applies to absolute, not relative, frequencies. For example, let us assume that f_o is 1280 acres and f is 640 acres, which is also equivalent in area to f_o of 2 square miles and f of one square mile. However, chi square derived by using acres is

$$\frac{(1280 - 640)^2}{640} = 640$$

and that by using square miles is

$$\frac{(2 - 1)^2}{1} = 1.$$

Inasmuch as the chi square table would be entered with the same number of degrees of freedom irrespective of whether acres or square miles were the units, tests of significance would differ. It was this aspect of the use of unit areas of one acre that Reynolds seems to have referred to in discussing the dissertation upon which Zobler's present article was based.[3] Thus, if Zobler had collected his data upon the basis of square miles, rather than acres, and then applied the chi square test, the results would have differed from those obtained by using acres.

If Zobler could demonstrate that in his study there was a theoretical basis for the selection of unit areas of one acre, then the above criticism would obviously not apply, because the frequencies would be absolute, not relative. If, however, there is no theoretical basis for selecting a unit area of one acre, then the use of the chi square test seems invalid, and some other statistical technique should be used.

It should be stressed that the above objection is not meant to apply to all of Zobler's stimulating paper, but only to that portion involving comparisons of acres. As is well known, it is far easier to criticize, as this writer has done, than it is to undertake research, as Zobler has done. In this writer's view, Zobler should be congratulated for taking a forward step by using statistical techniques in regional analysis.

[1] Leonard Zobler, "Statistical Testing of Regional Boundaries," *Annals*, Association of American Geographers, Vol. 47 (1957), pp. 83–95.

[2] Frederick C. Mills, *Statistical Methods* (third ed.; Henry Holt & Co., 1955), p. 538.

[3] Robert B. Reynolds, "Statistical Methods in Geographical Research," *Geographical Review*, Vol. 46 (1956), p. 130.

JOSEPH BURTON KINCER, 1874–1954

JOSEPH BURTON KINCER, 1874–1954

S. S. VISHER
Indiana University

MR. KINCER was voted by the AAG Council as worthy of a memorial in the *Annals* because of his significant contributions to the study of climate of the United States. For four decades, while with the U.S. Weather Bureau, he analyzed with exceptional skill the climatological data gathered by the numerous cooperative and other observers, and presented his findings in many maps and diagrams. He was chief of the Bureau's Division of Agricultural Meteorology and its successor, the Division of Climate and Crop Weather, from 1925 to 1944 (when, at 70, he retired). He compiled the two chief sections on climate for the official *Atlas of American Agriculture* (1922, 1928, 1936). For *Climate and Man* (the 1941 Yearbook of the Department of Agriculture) he wrote the section "Climate and Weather Data for the United States," and prepared the 46 page-size maps on "The Climates of the United States." He also directed the compilation of the climatic summaries for individual states, which comprise the final nearly 500 pages of that volume. In the *Weekly Weather and Crop Bulletin*, which he edited for twenty years, he presented some novel maps and diagrams. Others were published in the official *Monthly Weather Review*, in special official bulletins, or on the backs of particular daily weather maps. Among non-governmental publications by him are an excellent study of "The Climate of the Great Plains," in the *Annals* of the AAG (1923), and "Evidences of Climatic Changes" in the *Transactions* of the American Geophysical Union (1946). The most widely quoted of his publications is *Is Our Climate Changing?* a bulletin issued by the Illinois Farmers Institute in 1937.

Mr. Kincer was born in Wythe County, Virginia, November 15, 1874, and died in Washington, D.C., December 14, 1954. He was educated in private schools. He taught public schools in his native county for seven years and served as deputy county treasurer there for three years. He entered the U.S. Weather Bureau in 1904 and was promoted successively until he became Principal Meteorologist and Chief of the Division of Climate and Crop Weather. He was elected president of the American Meteorological Society for 1936–37.

In the first six years of the decade between his retirement and his death at 80, he did much to assist me in the preparation of *Climatic Atlas of the United States*, which is dedicated particularly to him because his work and cooperation greatly facilitated its preparation. About a fifth of the nearly 1,000 maps of the *Atlas* are adaptations of maps of his, and many of the remainder were made possible by his work or influence in obtaining the unpublished data.

Mr. Kincer was interested especially in climatic facts and their mapping rather than in their interpretation. Hence he was a climatographer rather than a climatologist or meteorologist. He had special interest in averages and departures therefrom. Supplementing the climatic maps of *Climate and Man* and those in the *Atlas of American Agriculture* (edited by O. E. Baker), two special bulletins issued by the Weather Bureau may be mentioned as illustrations of this interest: *Normal Weather for the United States* (1943) and *Maps of Seasonal Precipitation* with tables for the wettest and driest seasons and years (1942). Another illustration is the development of "degree-days," a clever method of revealing the number of days with average temperatures below 65° and the amount of their coldness, and indirectly disclosing the wide regional variation and year-to-year fluctuation in these data.

Although Kincer's maps become out of date with the accumulation of additional records, his productive pioneer work in the study of the climate of the United States has won him an honored place in the roster of American scientists.

SIDMAN PARMELEE POOLE, 1893–1955

SIDMAN PARMELEE POOLE, 1893–1955

JOSEPH A. RUSSELL
University of Illinois

SIDMAN PARMELEE POOLE was born in Syracuse, New York, on October 19, 1893. Between that time and his death on October 28, 1955, he performed significant services to his country, his profession, and to Syracuse University and the University of Virginia.

Sidman Poole, more than any other person, made it possible for geographers to use their knowledge and techniques in defense of their country during World War II. The little-appreciated and virtually unexploited field of military geography which had been neglected since Douglas Johnson's "Battlefields of the First World War"[1] was re-created by Sidman Poole and those who worked with him in the Armed Forces and who followed him in other defense agencies.

At no other time in history had it been so important that major military decisions be predicated upon accurate, detailed information about the physical and human environments of the world. Yet, such knowledge was meager at all levels of military intelligence and America's geographers provided a potential source of much of the required environmental data. This presented a stupendous challenge, one made even more difficult because so few geographers had an understanding of military activities and needs. Sidman Parmelee Poole, whose love of geography had been kindled during military service in France during World War I, was one of the handful of geographers who had developed a continuing interest in the military implications of geography. He was called to active duty in the troubled days of 1940, before Pearl Harbor, and with singular purpose he helped to organize and became Chief of the Topographic Branch of G-2, which he commanded throughout the War.

Sidman Parmelee Poole dedicated his professional career to establishing geography as a significant tool in the formulation of military plans. That he was successful is attested by the regard with which the profession is now held, not only in intelligence operations and in broad research programs of the Department of Defense, but in the training of officers for all services. He has left his recommendations for the preparation of geographers for military duty in "The Training of Military Geographers," *Annals*, Association of American Geographers, 34 (1944), pp. 202–206. His contribution to his chosen profession has been a lasting one; in contrast to the decline in the appreciation of military geography that followed World War I, each of the defense forces of the United States, as well as civilian intelligence agencies, has maintained substantial interests in geographic research and applications.

Upon his return from duty in World War I, Sidman Parmelee Poole earned the B.S. and M.S. degrees from Syracuse University in 1921 and 1925. His Ph.D. was awarded by the University of Chicago in 1932. First appointed to the faculty of Syracuse University in 1921, he held the rank of Professor of Geography at the time he was called to Washington to mobilize the geographers for war service. Primary interests in Latin America and Europe led to participation in field research in the Andes in 1930–31, Brittany in 1933, and Yucatan in 1937–38.

After having been decorated by the governments of both the United States and Great Britain for his contribution to the winning of World War II, Sidman Parmelee Poole, in 1946, became Professor of Geography and Chairman of the School of Geography at the University of Virginia. Here he continued his devotion to the use of geography in contributing to the solutions of military problems. The Virginia Geographical Institute, which he founded and directed, provided significant research for the Army Corps of Engineers, the Navy Hydrographic Office, and the Office of Naval Research.

Despite his major contribution to geography as a military geographer, Sidman Parmelee Poole will be best remembered for his powerful personality: "Sidman Poole was a man of great personal charm. His intellectual curiosity extended far beyond the bonds of his professional interests, and his mental versatility rendered him an engaging conversa-

[1] D. W. Johnson, *Battlefields of the World War*, American Geographical Society, Research Series, No. 3, New York, 1921.

tionalist. Nothing human he considered alien to himself, and this bond of human sympathy brought him a host of friends in all walks of life. His personal qualities of unfailing cheerfulness and sincerity, his exuberant vitality, his zest for living and learning, endeared him to students in and out of the classroom, and won for him the esteem and affection of his colleagues. His ever-ready but kindly wit, his talent for friendship, his fine camaraderie and vivacity in his social contacts with his colleagues make his going a sore loss to the community of the University."[2]

[2] From a resolution adopted by the Faculty of the College of Arts and Sciences of the University of Virginia, February 29, 1956.

MALCOLM JARVIS PROUDFOOT, 1907–1955

MALCOLM JARVIS PROUDFOOT, 1907–1955

CLARENCE E. BATSCHELET
United States Bureau of the Census

MALCOLM JARVIS PROUDFOOT, associate professor of geography at Northwestern University, died suddenly on November 21, 1955, in Oxford, England, at the age of 48. He was on a year's leave of absence as a Guggenheim Fellow, engaged in research on wartime and postwar migration into the United Kingdom. Geography has lost a dynamic and inspiring disciple, and colleagues and students have lost a penetrating critic, sincere friend, and keen advisor.

His short career was colorful and yet at all times demonstrated a dynamic personality with unlimited energy. His interest in urban geography manifested itself early while he was a student of geography at the University of Chicago, from which school he received his Bachelor of Philosophy degree in 1928, Master of Science in 1930, and Doctor of Philosophy in Geography in 1936.

From August 1934 to November 1935 he served as associate geographer with the Tennessee Valley Authority. Here he was directly interested in land classification and land use mapping, contributing to the development and application of field techniques, and was the collaborator with G. D. Hudson in two important publications on land classification and use.

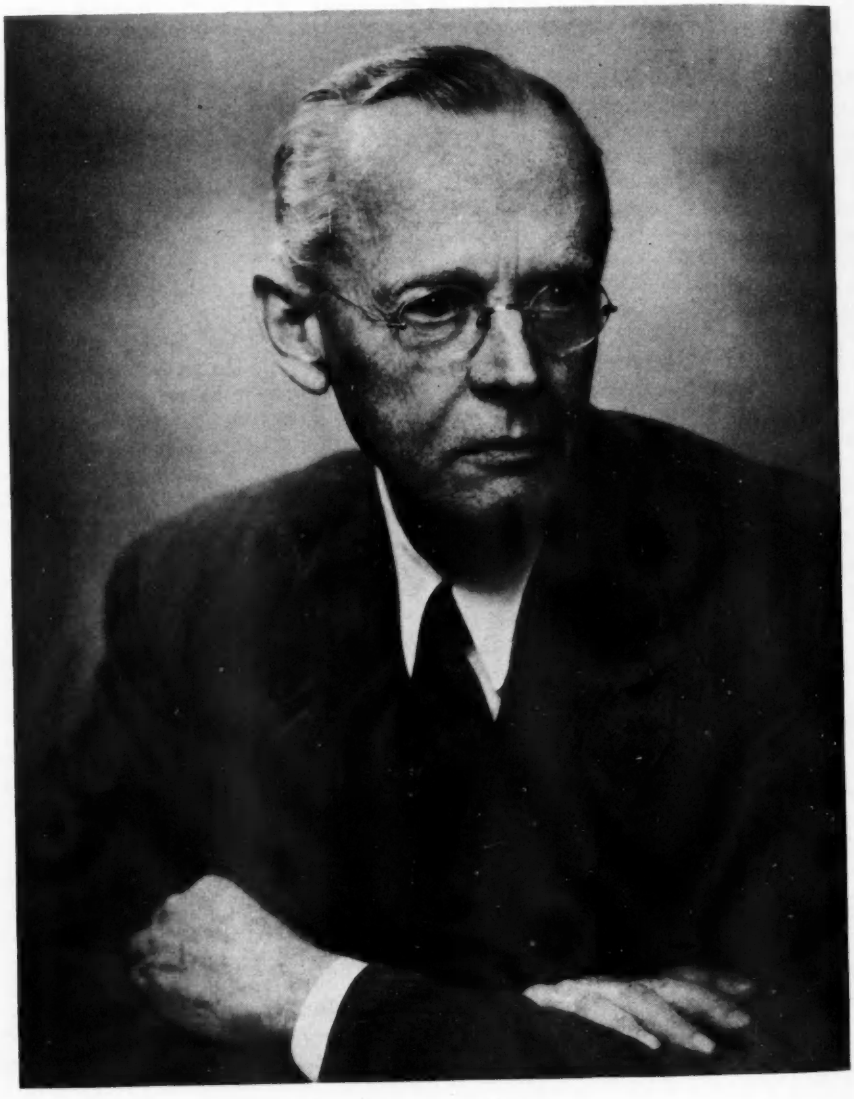
He came to the Bureau of the Census in November 1935 and the following year completed his doctoral dissertation, "The Major Outlying Business Centers of Chicago." This publication is regarded by many urban geographers as an outstanding analysis of commercial structures. For the Business Census of 1935 he analyzed the "Intra-City Business Census Statistics for Philadelphia," one of the first studies of this type to be compiled. He not only prepared the geographic manual but took an active and inspiring part in the training of personnel for the 1940 Census. While he was in the Bureau serving as assistant geographer he began his book, "Measurement of Geographic Areas," for which he was awarded a prize from the Geographical Society in Chicago in 1948.

His work in the Bureau of the Census was interrupted while he served with the rank of Lieutenant Colonel in the military government, in the Division of Supreme Headquarters, Allied Expeditionary Force. Here his duties involved working with refugees and displaced persons. His contributions were of such an outstanding character that he was twice decorated by the United States Army and by each of four Allied governments.

After the War he returned to his former position as assistant director of the Bureau of the Census, which position he had acquired in 1941. His sojourn lasted only a few months as he left the same year (November 1946) to accept a professorship at Northwestern University. While in the Bureau he had a yen to teach as he felt that he was especially qualified and experienced to teach geography.

His interest in Europe and his extensive travels permitted him to carry on extensive research and field work, and as a result he was recognized as an outstanding authority in geographic work.

He attended many international population conferences and geographic meetings presenting papers and actively participating in their discussions. He was an active member of the leading geographical societies in the United States and Europe and his contacts were numerous. He frequently represented societies and other organizations at these international conferences. His rich experience, breadth of interests, keen insight, and challenging ideas created among publishers a desire that he review outstanding publications. He possessed a talent which unfortunately too few scholars acquire and he will long be remembered by his fellow workers and students for his keen observations, his forceful expressions, and careful workmanship, which created within those with whom he associated an inspiration to achieve the best.



JOHN LYON RICH, 1884–1956

JOHN LYON RICH, 1884–1956

GEORGE B. BARBOUR
University of Cincinnati

JOHN LYON RICH was born at Hobart, New York, on December 1, 1884, and grew up on a dairy farm in that agricultural corner of the Catskills, to which the Rich and Lyon families had moved from nearer the coast three generations earlier. He attended the local school but seized every chance to be out of doors. When not in class or engaged in the exciting game of "beam tag"—played along the bare timbers of a barn across the road from the one-room schoolhouse—he was sure to be off with rod or gun over the countryside until, as he himself recalled with a chuckle, a neighbor remarked, "That John Rich won't ever amount to anything. He spends all his time hunting woodchucks." Though Mr. Rich did not want his son to be "tied to a cow's tail," a college training seemed out of reach. But John's interest in nature, his persistent curiosity, and a shelf of serious books accumulated over the generations augured for a scientific career. The boy would bury his head for hours in the volumes of *Popular Philosophy, or the Book of Nature Laid Open—with suitable Moral Reflections calculated to excite Devotional Feeling in the Breasts of the Young* printed at Haddington, Scotland, in 1826. Fully half the pages of this older-day "outline of science" are given to a very creditable survey of geological ideas then understood. When he found a copy of Laconte on the school bookshelf, the die was cast.

An uncle encouraged him to enter the district competitive scholarship examinations for Cornell. At sixteen, John was probably the youngest candidate. No one was more surprised than he when his name appeared at the head of the successful list. In his second year at Ithaca, his inquiring mind caught the attention of Gilbert Harris and Ralph Tarr. For the summer of 1905 and half John's senior year thereafter, Harris took him as field assistant to map the Winnfield Quadrangle of Louisiana. Here he learned reconnaissance mapping, getting direction by compass, ruler and plane-table, and pacing distances between stations, aneroid in his pocket. The next summer Tarr took him to Alaska on the Malaspina expedition for the U.S. Geological Survey and the National Geographic Society. That same year, while he was still an undergraduate, brought another distinction, the publication of his first paper—on local glaciation in the Catskills—in the *Journal of Geology*, beside one on solifluction from the pen of the Swedish geologist, J. Gunar Andersson.

In the summer of 1907, he won his master's degree and then did field work for the U.S. Geological Survey under Schultz and Veatch on the coal resources of Wyoming. Two years later he worked on the southern Californian iron ores and in 1910 joined Sidney Paige's party, surveying the Silver City Quadrangle in New Mexico. He took his examination for the doctorate at Cornell in 1911. By this time he had half a dozen publications to his name and accepted an instructorship in geography and physiography at the University of Illinois, for which he prepared by a round of visits to geography departments at universities in France, Switzerland, Germany, and Great Britain.

Seven years later the country was at war, and Captain Rich was in uniform, assigned to the Graphic Section of the Military Intelligence Division of the U.S. Army. His greatest contribution lay in the reorganization of the General Staff Map Room at the Army War College. After the first World War he worked for three years with oil companies, following this by nine years as consulting petroleum geologist. While working for Argus Oil under Eliot Blackwelder, he successfully turned his physiographic training to the shoestring sand problem with which his name has since been associated. "Moving Underground Water as a Primary Cause of the Migration and Accumulation of Oil and Gas" appeared in *Economic Geology* in 1921, followed in 1923 and 1931 by further fundamental papers in the bulletin of the American Association of Petroleum Geologists. For four years he edited the bulletin, adding to his personal contacts and range of useful experiences.

In the field, his tall spare figure was always on the move, but his eye missed nothing. In addition to his keen powers of observation, critical deduction and methodical workman-

ship, Rich showed an amazing versatility. He was equally at home whether the discussion was on petroleum, economic or structural geology, geomorphology, photogrammetry, sedimentology, metamorphism, or orogenesis. He challenged accepted ideas in all these fields. As a result, his careful writings represent significant contributions to science, many of which have now gained acceptance. Of the 120 papers from his pen, 48 were concerned with geologic processes; 17 were on stratigraphy, structural geology and petrography; an equal number dealt with oil origin and accumulation; eight were on field geology, and the balance on geography, meteorology, educational philosophy, or miscellaneous topics.

When Charles Behre was called to Columbia, John Rich was Nevin Fenneman's natural choice as professor of economic geology at the University of Cincinnati. But nothing could repress Rich's constant interest in the adjustment of man to his physical environment. Besides, Fenneman himself was a geographer and physiographer at heart. The acclaim given to a *Geographical Review* paper based on air photos taken when Dr. Rich flew across the Appalachians between Cincinnati and Washington led to the American Geographical Society's interest in a similar analysis of air photographs taken across "The Face of South America" (1942). The meticulous care with which he recorded every Leica exposure, and the thoroughness with which the exact position and field subtended by each frame was checked and rechecked before being transferred to the base maps in that monograph, were typical of the high standards he demanded of himself and others. He was dramatically impatient at sloppy or superficial work, or at laziness in thought or action. But his face lit up when his eye fell on sound work carefully presented.

On Fenneman's retirement Bucher replaced him as head of the Department of Geography and Geology, but within three years followed Behre to Columbia. John Rich was the obvious successor to the headship.

During the Cincinnati years, his studies were varied—the Cumberland Thrust Block (1933), the evolution of rock fans and pediments (1935), and multiple erosion surfaces (1938).

During the Second World War, Rich was called to Washington as technical consultant to the Petroleum Administration for War, and flew across the Atlantic, North Africa, and Southern Asia to join the American War Production Mission at Chungking from which he visited the oil fields of Northwest China. His assignment was to determine, in cooperation with the Chinese geologists, the adequacy of the oil reserve to support a combined attack from the northwest on the Japanese army of occupation. His field colleague was E. M. Bien, who had worked with Barbour in Peking prior to the Japanese invasion. Dr. Bien wrote thus:

> Dr. Rich came to the Laochunmiao Oilfield in Kansu about three months before VJ-Day. I recall our first meeting. I saw a tall man with discerning and thoughtful eyes, dressed in army fatigues. My colleagues and I were immediately impressed. We were already familiar with his work on the shoestring sands of Kansas. At the same time, we were afraid a man of his standing might be too authoritative to permit an atmosphere favorable for smooth and close cooperation. Our misgivings were soon dispelled. Dr. Rich took the prescribed round of formal banquets in his stride without making a single *faux-pas* in the eyes of his hosts. In discussions in the field and in the office he quickly taught us his ideas on the hydraulic theory of oil migration and on the importance of aerial observation, especially in arid terrains like the north-west where structures are so well exposed at the surface. He was always ready to listen to other men's opinions. Our problem was to estimate the depth to the producing horizon prior to drilling. It invariably proved to be deeper than anticipated. In his final report he wrote with characteristic fairness: "This difference has attracted the attention of the local geologists who are inclined to attribute it to the thickening of the formations near the axis by inter-stratal slippage and semi-plastic injection into the axial part of the field. Their explanation seems probably to be correct." On the trip by truck to Huangshui west of Lanchow, Dr. Rich was indisposed. He needed a rest and we begged him not to strain himself by going out to the field. He finally consented to take a day off and accepted our field observations. On the flight back to Chungking from Chiuchuan, Dr. Rich noted a conspicuous structure north of Lanchow. He wrote me immediately about it, enclosing a sketch map. This prompted a ground survey by the geologic staff of the Kansu Petroleum Administration which completely corroborated Dr. Rich's interpretation based on air observation alone. We admired his quick grasp of the overall picture. But we never were able to fulfill his wish of making observations from the back of a camel! He was a man remembered by all of those who had the good fortune of coming into contact with him.

The first International Geological Congress after the War took Rich to London in 1948. An important result of a pre-congress trip was his paper on the flow-markings in the Silurian

marine strata of Wales (1950). A year later ideas presented there were expanded under the title *Three Critical Environments of Deposition* (1951).

In 1951, Rich was granted leave to go as Visiting Professor of Geology to the University of São Paulo. His ability to cope with the language led to rewarding field trips with his Brazilian colleagues.

Retirement at the age of 70 meant no cessation from active duty, and was only a chance to complete studies that had been pushed aside under the pressure of academic duties. One of these was a brilliant paper on the interpretation of landforms by aerial reconnaissance, which he prepared for the *Manual of Photo-interpretation* of the American Photogrammetric Society, under the title "Geomorphic Basis for Photogeology." Another dealt with the paleogeography of the Chattanooga–Ohio Black Shale epoch.

Despite a premonitory heart attack on the eve of departure for his second tour as A.A.-P.G. Distinguished Lecturer, his wife's constant care seemed to restore him to full health, and he was able to carry through the gruelling circuit without misadventure. It was an honor he cherished.

Since childhood Dr. Rich had held the conviction that geographic environment could be used more fully for the benefit of the people. It seemed like a dream come true when Governor Lausche appointed him to the State of Ohio's Natural Resources Commission. He devoted himself wholeheartedly to its meetings and inspection trips to forests, state parks, subterranean salt deposits, and water-supply areas. Death when it came found him as he would have wished—in harness. Late in the day on May 17, 1956, the commission party pulled up to the cabins at Lake Hope State Park in heavy rain. Rich impulsively carried his bags from the car and then went to find logs for a fire against the inclement weather. But he had over-exerted himself and was taken to the local hospital at Logan, where he died four days later.

When presenting Dr. Rich for the distinction of honorary membership in the Association of American Petroleum Geologists in 1954 Morris M. Leighton, former chief of the Illinois Geological Survey, ended aptly by quoting Wordsworth's couplet,

> Enough if something from our hands have power
> To live, and act, and serve the future . . . . . .

REVIEW ARTICLES

WHITHER POLITICAL GEOGRAPHY?

Has there ever been a period in history other than the present when man has been so troubled about his behavior and the complex of motivations which drive him simultaneously toward great achievements and self-destruction? If man, as Aristotle wrote, is a political animal, then he is now being subjected to every available technique to determine how and why he behaves as he does, especially when engaged in political activities. The question concerns the geographer as well as historians and other social scientists. Indeed, in view of the rash of substantive works published recently in political geography, the geographer seems deeply interested. Such a manifestation of interest in a field which hitherto has tended to lag behind the other fields of geography calls for some discussion, particularly about the state of the field today, and the nature of the contribution geographers make to the over-all understanding of the basic problems confronting political man.

The phenomenon called the "state" has been accepted by geographers generally as the formal or central subject matter of political geography.[1] As a political–territorial unit, revealed on the ordinary political map and implanted in the consciousness of every school boy, the state represents the organization of a particular area of the surface of the earth and of the people who permanently reside therein.[2] Ordinarily, the state is conceived in terms of the "Nation State" which, emerging slowly and painfully in western Europe after the Reformation, is now being copied in most other parts of the world. The state, national or otherwise, however, is *also* an institution, involving a system of relationships which men have established among themselves in order to secure certain objectives and to carry on certain activities. There is, therefore, a dual aspect to the concept of statehood, suggesting, on the one hand, the organization of territory for specific political ends and, on the other, a system of relationships created among the inhabitants of that territory.

In studying the state, geographers have been aware of this dualism and, while proceeding to deal geographically—although perhaps not always to everyone's satisfaction—with the territorial unit, have often seemed perplexed as to how, if at all, they should treat the institutional relationships. As a result of this confusion, the development of political geography as a systematic field of study has tended to lag behind other branches of the discipline. Part of the difficulty, of course, may spring from what geographers have thought their field to be.

At the same time, the problems confronting the geographer are likely to reflect the situation inherent in political science, the related "discipline," where both method and subject matter are highly diverse.[3] What unity, one may ask, is there in a discipline which includes the study of the history of political thought, political philosophy, political institutions, political power, international relations, and the process of government at every level? Yet, according to Morgenthau, there *is* a common denominator which unites the field; it is a common subject matter defined in most general terms, an "orientation towards the nature and activities of the state and toward activities which have in turn a direct bearing upon the state."[4] Obviously, then, the major link between political science and political geography is a sharing of interest in the state.

As a social science, political science ultimately is focussed on man in society, organized for the purposes of promoting and carrying on political activities, of establishing a government, of creating a state; above all, the political scientist is concerned with the system of relationships referred to above. Political geography, as a branch of geography with its roots in the natural sciences as well, has concentrated on the territorial aspect: what territory, how organized, and, to some extent perhaps, why. However, the division of labor between the political geographer and the political scientist with respect to the state, as evidenced from past and present writing, may well be wholly unrealistic.

Some years ago, in his presidential address Hartshorne appealed for "a more geographical political geography";[5] yet, his "functional approach" to the subject demanded considerable knowledge of and insight into institutional relationships. On the other hand, Sprout had already suggested that "political geography should become an established political science field," since, as he believed, regional geography is the bedrock upon which political science is founded.[6] More recently, he and Margaret Sprout have in essence restated the thesis, indicating that political science, with respect to international relations, should become more geographical.[7] This restatement provides a convenient benchmark from

---

[1] Stephen B. Jones, who kindly read the manuscript, would object to putting the state in such a central position in political geography; his inclination would be to define political geography as the study of political activity in its areal context, with the state representing only one manifestation of political activity, though, to be sure, the focus of our most intense loyalties and strongest political control.

[2] Richard Hartshorne, "Political Geography," in *American Geography, Inventory and Prospect*, Preston E. James and Clarence F. Jones, eds. (Syracuse: Syracuse University Press, 1954), pp. 169–225.

[3] Thomas I. Cook, "The Methods of Political Science, Chiefly in the United States," in *Contemporary Political Science: A Survey of Methods, Research and Teaching* (Paris: UNESCO, 1950), pp. 75–90.

[4] Hans J. Morgenthau, "Reflections on the State of Political Science," *The Review of Politics*, Vol. 17 (October, 1955), pp. 431–60.

[5] Richard Hartshorne, "The Functional Approach in Political Geography," *Annals*, Association of American Geographers, Vol. XL (June, 1950), pp. 95–130.

[6] Harold H. Sprout, "Political Geography as a Political Science Field," *The American Political Science Review*, Vol. XXV (June, 1931), pp. 439–42.

[7] Harold and Margaret Sprout, *Man–Milieu Relationship Hypotheses in the Context of International Relations*. Princeton: Princeton University Center of International Studies, 1956. 101 pp. n.p.

which to embark on a review of the recent substantive works in political geography.

The objective of the Sprouts's monograph is to "examine critically and comparatively general hypotheses regarding the relations between human affairs, on the one hand, and the environment, or milieu, in which these occur." According to the authors, an examination of these relationships has been significantly absent from the modern literature of political theory. Such a review, therefore, is designed to reveal "the utilities and limitations of these hypotheses as tools for calculating the capabilities of states, and for explaining and predicting the actions and reactions of states, within the frame of reference called international politics." Because of its implications, the study has a place in a review of recent literature dealing with political geography and the state.

After carefully defining their terms—the word "milieu" is used to refer to the "totality of environing factors rather than to some sub-category thereof"—the authors proceed to an evaluation of the hypotheses which they feel have a bearing on man–milieu relationships, namely: environmental determinism, mild environmentalism, environmental possibilism, environmental probabilism, cognitive behaviorism, and other more general theories of explanation and prediction.

One may question the need for and value of a re-examination of determinism or of mild environmentalism. Undoubtedly, there may still exist an identification among some political scientists of geographic study with environmentalism, but the latter no longer constitutes the organizing principle of geographic theory. However, the survey does provide the Sprouts with an opportunity to suggest that the reaction to determinism which occurred several decades ago may have been a case of "the proverbial throwing out the baby with the bath." Man's non-volitional responses to environmental facts may set limits to human choice as well as to human capabilities.

The chapter on possibilism is, perhaps, the high point of their essay. Indeed, several of the authors of recent texts in political geography might well have benefited by the Sprouts's assessment had it been available to them. Possibilist analysis, of course, is conceived in terms of opportunities present in the milieu, but latent until man makes a decision and embarks upon a course of action toward some chosen end. Possibilism, however, provides no framework within which to determine how choices are made: motives and decisions are taken as given, not phenomena to be analyzed, explained, or predicted. Therefore, the possibilist in calculating state possibilities is compelled to proceed from some assumption of what man is attempting or will attempt to do.

Cognitive behaviorism, the awareness that man may react to the milieu not in terms of what the milieu is, but as it is conceived to be, provides a means of getting at the process of how choices are made. Yet, the hypothesis does have its limitations. While there may or may not be a maladjustment, more or less significant, between the milieu as perceived and the milieu as it is, there must be, as the authors point out, some raw material with which to work if state policy is to be successfully implemented. What a people think of themselves and more especially what the persons in control think of their state with respect to other states may have a very decided bearing on the formulation and subsequent implementation of policy. Nevertheless, a state is not a great industrial power, to cite an example (no matter what its citizens think or feel), unless it possesses or has access to the necessary resources and the means and ability to use them.

What is termed "environmental probabilism," involving the development and use of generalized models of man's reaction to a given milieu, the Sprouts find unacceptable for the study of international relations. Obviously the danger in constructing models is that the underlying assumptions may be faulty and tend to oversimplify the objectives of statecraft, especially in policy-making. To quote the Sprouts: "The analyst may easily overestimate the decision-maker's probable knowledge and understanding of the limitations implicit in his milieu. Or the analyst's own knowledge of that milieu may be seriously defective. Or he may assume greater rationality than actually prevails in the practice of statecraft. He may underestimate or overestimate the probable willingness of decision-makers to take risks. The analyst is especially prone to minimize or ignore differences in intellectual operations carried on in a language and in a social context radically different from his own. He may unconsciously ascribe to foreign decision-makers his own set of values and his own standards of prudent conduct."

It is questionable, however, if the Sprouts's general argument will find acceptance among most political scientists.[8] The implications of probabilism have already created among them issues which have been debated at length, frequently with considerable heat. In this context, it is interesting to note that Isard has urged that political science train analysts capable of undertaking what is described as "sophisticated research," but at the same time he has stated that "no theoretical model or advanced statistical procedure can substitute for political acumen or for balanced judgment and intuition which are basic to policy decisions."[9]

In concluding, the Sprouts suggest that the most suitable frame of reference for the study of international relations lies within possibilism linked to cognitive behaviorism.

The several substantive works in political geography published in recent years confirm that the unit area of study is primarily the state, although there may be differences of opinion as to its nature and *raison d'être*. Generally, the subject matter is treated within the framework of possibilism with which a cognitive behaviorism is blended. Take, for example, the work of Weigert and his five associates.[10]

Weigert, the author of the introductory sections of the book, notes "the healthy reaction" that occurred

---

[8] See: Henry S. Kariel, "Political Science in the United States: Reflections on One of its Trends," *Political Studies*, Vol. IV (June, 1956), pp. 113–27; Cook, *loc. cit.*

[9] Walter Isard, "Potential Contributions of Regional Science to the Field of Political Science," *Papers and Proceedings*, The Regional Science Association, Vol. 3 (1957), pp. 29–32.

[10] Hans W. Weigert and others, *Principles of Political Geography*. New York: Appleton–Century–Crofts, Inc., 1957. viii and 723 pp. Maps, index. $7.95. The five coauthors are Henry Brodie, Edward W. Doherty, John R. Fernstrom, Eric Fischer, and Dudley Kirk. Of these, Weigert, Fischer, and Fernstrom are professional geographers, the latter undertaking the drafting of the maps.

in the United States against the theories of determinism and of possibilism. However, his interpretation of political geography seems to involve possibilism or a modified version of environmentalism. Political geography, the text reads, is "a legitimate child of human geography. Both deal with the interplay of physical and human factors, with the interrelationship between earth and man. Both try to discover and explain the influences of the physical world on human society and the limitations it puts on human activities; they deal with the diverse manifestations of a symbiosis of nature and man." Further, "it is, perhaps, the roots of the human groups in their natural environment that most influence their development. These are, however, not the only formative factors in human society. Historical and sociological motivations, as well as cultural influences, cannot be discounted. Yet, to be rooted in a natural and cultural landscape and environment is the essence of life to the individual and to the group . . . ." At the same time, a quotation from James introduces another concept which suggests cognitive behaviorism. James states that "the physical character of the earth has different meaning for different people: that the significance to man of the physical environment is a function of the attitudes, objectives, and technical abilities of man himself. With each change in any of the elements of the human culture the resource base provided by the earth must be re-evaluated."[11] However, it is questionable whether the authors of the text clearly appreciate the implications of the James statement.

Throughout two-thirds of the book, a so-called "functional approach" is adopted. However, the approach is not Hartshorne's, nor is it clearly defined. One wonders if it might not better be called a "systematic topical approach." Yet, among the topics considered nowhere is there a discussion of the attitudes and objectives of man with respect to the organization of territory for political purposes, although it might be argued that linguistics and religion, which are referred to, have some bearing on the matter.

While topical treatment is given to the spatial and to the human and cultural factors in political geography, the economic factor is discussed within a regional framework. Indeed, the latter third of the book reads like regional economic geography. Here, to a considerable degree, the economic factor is thought of with respect to the power base of the state. Power, of course, is not the only motivation of states, and certainly an assessment of state power should involve much more than the economic factor. Clearly the weakness of the book lies in its attempt to work with two approaches which destroy its continuity and prevent integration of the many factors involved.

Since the second edition of *Elements of Political Geography* by Van Valkenburg and Stotz also is topically organized, it seems pertinent to refer to it here, although in publication it antedates the Weigert book by several years.[12] In viewpoint, it is more consistent than the latter, although somewhat more elementary. Considerable emphasis, as in the first edition, is placed on the physical elements which are tediously enumerated. However, the section on the economic elements has been expanded to advantage, and the human–cultural elements have received fuller treatment. Major criticisms that may be levelled at Van Valkenburg and Stotz involve their failure to provide insights into political processes and motivations and to integrate, in conclusion, the various elements of political geography. It is, in short, a "static" book.

The second edition of *World Political Geography* by Pearcy and some 26 associates[13] differs notably from that of Weigert and his associates not only in organization and presentation of material, but also in underlying concepts with respect to the nature of the state. Political geography "consists of the description and analysis of the politically organized area," but it also is related to the broad field of power politics. Thus, the central theme around which the discussion of the world political pattern is woven is power, with power potential conceived as the ability of the state to determine and implement policy at any level.

Yet, in the international sphere there are obviously two levels of power, if not more. "Great" powers have always exerted, or attempted to exert, a "primacy" over the smaller and weaker. The somewhat stock, tediously factual, continent-by-continent and country-by-country approach of the book detracts from these political realities. Moreover, if in the realm of power politics, power potential (or, in the last analysis, war potential) is to be studied and analyzed with respect to the sovereign states, large or small, greater penetration of subject matter than is displayed by the authors, many of whom doubtlessly do not subscribe to the power thesis anyway, is required. Above all, the role of power in the present age must be placed in perspective—perhaps somewhat along the lines developed by Knorr in his systematic economic study, *The War Potential of Nations*[14]—if a political geographical study of state power is to have any value whatever.

The emphasis on power, pronounced in the introductory chapters of Pearcy's book, is missing from that of Alexander.[15] The latter states simply that he intends to "analyze the complex pattern of political units throughout the world. On the global or continental scale consideration is given to such forces as nationalism, imperialism, and economic competition, as they influence the political control of territory. On a more restricted scale national and local aspects of the political pattern are studied." Unlike most other American texts, Alexander's book enjoys the advantage, as being the work of one author, of maintaining a fairly consistent viewpoint throughout.

Insofar as political geography is concerned, its

---

[11] *Ibid.*, pp. 4, 5.

[12] Samuel Van Valkenburg and Carl L. Stotz, *Elements of Political Geography*. New York: Prentice–Hall, Inc., 1954. xvi and 400 pp. Maps, tables, charts, bibliography, index. $6.00.

[13] G. Etzel Pearcy and Associates, *World Political Geography*. New York: Thomas Y. Crowell Co., 1957. 2nd ed., xviii and 734 pp. Maps, index. $7.50.

[14] Klaus Knorr, *The War Potential of Nations* (Princeton, New Jersey: Princeton University Press, 1956). Not only does the author examine the nature of military power and the meaning and significance of war potential, but he also discusses at some length the "will to fight" and the administrative capacity for war, in addition to the economic capacity for war which constitutes the major section of the study.

[15] Lewis M. Alexander, *World Political Patterns*. Chicago: Rand McNally & Co., 1957. xii and 516 pp. Maps, tables, index. $6.00.

sphere of operation, according to Alexander, seems sharply delimited. Political geographers, he states on page 32, "are not concerned with the form or structure of government by itself, that is, the division of functions among executive, legislative, and judicial branches or whether it is democratic or totalitarian. These are matters for the political scientist." Unquestionably, these *are* matters for the political scientist, but the implications of a system of government or a political philosophy or ideology (sometimes expressed in the state's constitution) may have considerable geographical import and necessitate careful study on the part of the political geographer. Moreover, the structure and nature of a political system cannot be so dismissed. Not only may the system reveal something of the goals of the society involved, but it may provide insights into the character of the internal territorial organization as well as the direction which the state's international ambitions, if any, might take. This is not clearly brought out in the text. One wonders, too, why such forces as socialism and state capitalism—above all, the apparent drift toward the all-powerful regulatory state—are neglected as forces influencing the political control of area. Similarly, developments in Korea, Indochina, and perhaps in eastern Europe, as well as in the USSR and Communist China, cannot be thoroughly perceived without some consideration of the implications of Marxism, modified by Lenin and Stalin and other similar revolutionaries.

Some of the terminology as well as a number of statements throughout the text give cause for complaint. It is unfortunate that the author selects the term "responsible government" to refer to one of the five basic elements of the political region. Not only does it, among the member states of the Commonwealth, imply cabinet, and thus "responsible," government, but it also suggests by definition an accountability which such governments as those which exist behind the so-called Iron and Bamboo Curtains, and undoubtedly elsewhere, do not appear to possess. It is disturbing, too, that those member states of the Commonwealth which have retained the monarchical principle in their internal organization should continue to be referred to as "dominions," a usage which in recent years has fallen into abeyance. Canada, moreover, has legally dropped the title from its official name. Finally, the author seems to recognize only a saltwater imperialism, for the states of western Europe are discussed under the heading, "The Colonial Powers of Europe," whereas Soviet Russian imperialism is cloaked beneath the chapter entitled simply "The Soviet Union."

In some ways, *The Changing World* by East and Moodie,[16] with some eighteen collaborators, is a more satisfying book than either the Pearcy or the Alexander texts. Perhaps it is because the editors intend it to be no more than a geography of national and international affairs. Political geography as a field of study is not defined in so many words, and thus there is little or no formal methodological contribution. The only limitation placed on the twenty authors was that of length of text. Like the Pearcy work, it treats of national or state power and examines a host of pertinent problems, but this is attempted without resort to standard or standardized regional description. In effect, each chapter reads like a separate essay. Hence, as a survey of current affairs, with at the same time a strong historical flavor, it should rank high on the list, providing the student with a valuable source of information, intelligently handled.

Whenever geographers attempt an assessment of Soviet strength, it is often done with reference to Mackinder's "Heartland" theory. A chapter by East, entitled "The Soviet Union and the 'Heartland'," reflects the compulsion to test once again Mackinder's prognostications. The question that should be raised with respect to this habit and perhaps disposed of before long is simply this: Is Mackinder's hypothesis relevant to an evaluation of the Soviet position in the modern world and should it be thought of as a possible guide to future developments? Writes East: "Since the Soviet Union now controls almost all of the Heartland of Mackinder's conception, his proposition, translated into contemporary terms, invites *as a matter of no mere academic importance* [italics mine], the inquiry: how strong is the Soviet Union? . . . While it must be admitted that Mackinder's geopolitical philosophy has not become irrelevant to the international world of to-day, it is nevertheless clear that it contains generalizations and assumptions which scarcely withstand close analysis. And if, as must be assumed, one of the possibilities of the near future is the renewal of war on a grand scale, it is well to re-examine his sweeping prognostications to determine their present worth." If this is a matter of no mere academic importance, can it be disposed of within a few pages of text? If this problem is to be given the serious attention that it warrants, geographers might do better to examine carefully the writings of Lenin and his successors. Soviet geographers for their part seem to have given little attention to Mackinder (nor, it should be noted, to political geography as a field of study partly because it has been equated with geopolitics and held in low esteem). At any rate, Soviet strategists in planning the enhancement of Soviet power and prestige will certainly be prompted by the dynamics of their own system and guided by the unfolding of their own ideology, though modified to meet current needs and demands. There is, therefore, real danger in frequent re-statement of the "Heartland" hypothesis with respect to the Soviet Union basically because, as possibly the wrong frame of reference, it may conceal realities. Nevertheless, East's conclusions are reassuring for the moment. While the "Heartland" concept "takes note of certain geographical advantages of the USSR in warfare," he concedes, it "does not justify Mackinder's implied prediction about the future mastery of the world."

Goblet[17] and Maull[18] in their books lead to a somewhat different interpretation of the state in political geography. Both conceive of the state as an organism, and although the emphasis is slightly different, the general conclusions are similar.

[16] W. Gordon East and A. E. Moodie, eds., *The Changing World*. Yonkers-on-Hudson: World Book Company, 1956. xvii and 1040 pp. Maps, diagrams, index. $7.75.

[17] Y. M. Goblet, *Political Geography and the World Map*. New York: Frederick A. Praeger, 1955. xviii and 291 pp. Maps, illustrations. $6.50.

[18] Otto Maull, *Politische Geographie*. Berlin: Safari-Verlag, 1956. 624 pp. Maps, illustrations, index. $4.71.

"It is the task of the political geographer," writes Goblet, "to study . . . the territorial elements and complexes and the geographical forces and relationships which are at work in the formation, evolution and disintegration of most political organisms, notably of States." The sphere of political geography, therefore, involves the interactions between two basic elements, territory and man; out of this interaction is born an organic whole, real and living, which in time will wither away. Hence, the state is a man-created earth-bound being, endowed with personality which, in part, is a function of a particular geographical environment. The concept of the state as an organism is crudely conceived and hearkens back to the Darwinian process of evolution; few geographers in America will find themselves in sympathy with the author's point of view. Moreover, not only does he claim for political geography the ability to make "a critical geographical analysis of existing States . . . ," which is not satisfactorily realized in his own text, but he also suggests that the study "prepares schemes for the delimitation of the optimum territorial extent of each State. . . . In this way," he continues, "it may be regarded as an 'information service' for governments and their diplomatic corps." Unfortunately, the curtain dividing applied political geography and geopolitics has seldom been preserved by national officialdom.

Maull's work is considerably more stimulating than that of Goblet, if only because his concept of the state as an organism is more fully developed. Two ideas are basic to Maull: *Raumwesen* and *Raumorganismus*. *Raumwesen* constitutes the liaison between men as an organic group and territory, while *Raumorganismus* refers to the conscious and intense feelings which are generated in this relationship. Thus, the state is an organism sprung from a cell. These organisms, however, are not identical with biological organisms, and they have their own laws of development. They want to grow and expand, and it is in this process that they come in conflict with other states. This, then, is the cause for expansion and imperialism, a stage in development through which all states must pass. In view of recent history, Maull's interpretation sounds like an apology for Germany's behavior two decades ago.

Maull's concept of the state, which draws heavily from history for illustration, involves the Darwinian struggle for survival. It is reminiscent of Kjellen and reflects a philosophy which, in contributing to German *geopolitik* before World War II and to militant nationalism, is surely out of date in the atomic age. Moreover, the concept of the state as an organism is inherently inconsistent. Not only does the state not have Maull's "spiritual character," but as organized human society it obviously lacks the vital unity of the natural organism. Indeed, if human society were a generated organism, it would not require a government and there would be no state.

Goblet's classification of states into three categories, intensive, extensive, and mixed, represents his major contribution; otherwise, his study, which is systematic in approach, offers little that is new. The major portion of Maull's work, following the statement of his thesis, is regional, with brief historical, political, and cultural descriptions of the earth's political units.

Gottmann's systematic study of the state[19] has been reviewed briefly elsewhere,[20] and perhaps little can be added that has not already been said. Certainly the ideas which Gottmann generates, involving circulation or the "movement factor" and iconography or the system of symbols in which people believe,[21] are provocative and open up new vistas compelling further attention and study. They suggest a "dynamic" which Hartshorne was seeking to reach with his concept of centrifugal and centripetal forces. Gottmann's circulation, moreover, represents one of the links in Jones' unified field theory, though called "movement" by the latter, between the political idea and its ultimate manifestation in the political unit.[22] Both Gottmann and Jones are interested, in short, in interaction, which theoretically unites them with other geographers, similarly inclined but working in other fields of geography.

Interaction is expressed in somewhat different terms by Deutsch, to whom Jones has occasion to refer. Deutsch, in a thought-provoking book entitled *Nationalism and Social Communication*,[23] is seeking a method of investigation which will permit him to understand more clearly many of the intangibles associated with problems which such a phenomenon as nationalism creates for the social scientist. He is, in effect, concerned with developing a set of concepts other than that which, derived from mechanical equilibrium theories, "had little or no room for irreversible changes in structure, for growth, for self-transforming processes, in short, for history." In his idea of "complementarity of social communication," aimed at measuring communicative efficiency between groups, he is pointing toward a standardized method of analyzing Gottmann's iconography and circulation with respect to the group, a people, a nation, and/or an international organization, an effort which may have considerable value for the geographer, political or otherwise.

What conclusions can be drawn from this array of recent substantive works in political geography? What are the weaknesses in the field, and how may some of these be remedied in the future?

There is real need for the development of systematic thinking in political geography. The elements and principles which have been described and analyzed in recent systematic texts are sometimes trivial, disconnected, and undynamic. The so-called human element receives but superficial treatment, and attitudes and motivations, with respect to politics, are virtually ignored. Lip service is paid to the close

[19] Jean Gottmann, *La Politique des Etats et leur Geographie*. Paris: Librairie Armand Colin, 1952. xi and 228 pp. 670 fr.

[20] *Economic Geography*, Vol. 29 (January, 1953), p. 92; *The Geographical Review*, Vol. XLIII (April, 1953), p. 300.

[21] These ideas are also stated in: Gottmann, "Geography and International Relations," *World Politics*, Vol. III (January, 1951), pp. 153–73; "The Political Partitioning of Our World: An Attempt at Analysis," *World Politics*, Vol. IV (July, 1952), pp. 512–18.

[22] Stephen B. Jones, "A Unified Field Theory of Political Geography," *Annals*, Association of American Geographers, Vol. XLIV (June, 1954), pp. 111–23.

[23] Karl W. Deutsch, *Nationalism and Social Communication*. New York: John Wiley & Sons, Inc.; Cambridge: The Technology Press of the Massachusetts Institute of Technology, 1953. 292 pp.

relationship between political geography and political science, but fresh insights into political reality might be gained if geographers would sincerely study political science — and perhaps social psychology. The Sprout monograph, for example, offers certain hypotheses which might be effectively tested and developed by geographers in substantive writing. One wishes, too, that the concepts developed by Gottmann and Jones, with regard to circulation or movement, might be more fully demonstrated, especially at the local or domestic level which geographers have ignored while training their sights on the loftier international sphere.

The regional texts, which employ the country-by-country approach, also leave much to be desired. Not only do they contribute little methodologically to the field of political geography, but, by working the same materials over and over again, seem incapable of affording a fresh evaluation of modern political geographical problems. It may be that the authors, either singly or jointly, have attempted a much too ambitious project in a survey of the world, prompted, one suspects, by the commercial publishing houses. If the regional approach is to be employed, and its validity is not in question here, then it might be wise to restrict coverage to specific or limited areas in the study of which concepts might be applied, developed, altered, or discarded.

W. A. DOUGLAS JACKSON
University of Washington